

Not the percentage but the cash

How much should a nation spend on research and development? Advanced industrial countries generally devote between one and two per cent of their gross domestic product to R & D; usually somewhat more than half of the figure will come from government sources, the rest from private investment. Recent OECD figures (*OECD Observer* No. 97) show that Britain, devoting 2.1% of its GDP to R & D, still seems satisfyingly committed to science and technology. The reality is somewhat different.

Britain's GDP is (on a per capita basis) little more than half that of many other industrial nations. So if we ask how much money per capita is spent on R & D, Britain's position looks decidedly less distinguished. Whilst the US spends \$200 per capita annually, Germany \$176, Sweden \$171, the Netherlands \$146, France \$130 and Japan \$103, Britain only manages \$92. Comparisons look even less happy when an allowance is made for defence R & D in which the United Kingdom is a heavier spender than almost any other nation. The conclusion has to be that in civil

R & D expenditure the UK is quietly losing touch with the rest of the industrialised world; not through any visible lack of enthusiasm for science and technology but simply through lack of resources.

Talks of cuts in public expenditure are very much in the air at present as a new Conservative government tries to live up to its election pledge of tax reductions. On the other hand, the Advisory Board for the Research Councils recently made a strong plea for 4% real growth in science expenditure in the coming years. The many government departments with significant commitment to R & D will find it easiest to go for a middle path—no real growth. But in another ten years our relative scientific position will then be much worse and scientists will be taking every opportunity to emigrate.

There is no guarantee that a stronger commitment to science and technology will boost the economy—but there hardly seems any doubt that support which is relatively weaker on an international basis will do permanent harm to the nation. □

More on your fuel bill

THE winter in Britain is now over. Any further freak snow-storms, freezes, high winds and flooding should be attributed to next winter; this past winter, by general consensus a miserable one, cannot be blamed indefinitely. There is one final reckoning, however: the last quarter's gas and electricity bills sitting unwanted on doormats around the country.

These bills, everyone must know, are produced by computer. How else would we have a line reading 'VAT at 0% on £72.18=£0.00' or a 52 digit code number sternly marked 'for machine use only'? But what is surprising is not what the computer does to embellish these straightforward demands for money—it is what it leaves out.

Every organisation devoted to the selling of energy is at least nominally also interested in conserving the stuff. Conservation, however, is not just a question of double-glazing and thick sweaters; it is, or at least it should be, yet another part of that monster of the 1970s—the information industry.

The most valuable piece of information that any energy consumer can have at present is how conservation measures have cut down fuel bills. Did recently installed insulation in the attic have any perceptible effect on the

use of fuel? If so, that little bit of information, passed round from person to person, is a much greater incentive to others to invest in conservation measures than is any advertisement proclaiming the merits of insulation. If on the other hand fuel is not conserved thereby, then the consumer should equally be alerted that all is not well.

Presumably within the computers of the utilities there is information on each consumer's energy usage over a period of many years. This by itself would not be adequate to allow for weather fluctuations. But someone, somewhere must be collecting rudimentary data on how cold it is—such as the simple but valuable 'degree-days' measure widely publicised in the United States. It is surely not beyond the wit of the energy utilities to provide a line on each bill giving, say, the number of units consumed per degree-day for the quarter in question—and corresponding figures for the past five years.

No doubt there can be endless discussion about the merits of various units, the possibility that consumers will misinterpret the figures, the fraction of fuel that is used for heating and so on. But anything would be better than today's total lack of information. And there is room on the bill—that line about $0 \times £72.18 = 0$ could go for a start. □



Problems with the shuttle cause concern for US space science programmes

David Dickson reports from Washington

TECHNICAL problems and the related cost over-runs of the National Aeronautics and Space Administration's space shuttle programme are raising serious concerns about the possible impact on planned space science missions. So far, no major changes in these missions have been necessary. But with performance margins being squeezed increasingly tightly, contingency plans are already being drawn up for two planetary missions that depend on space shuttle launches—one to send a pair of space vehicles to Jupiter, and the other to send two vehicles to the Sun—in case any further difficulties arise.

Equally important in the medium term is the effect the current shuttle problems could have on new space science programmes. At present Congress seems prepared to accept that the military need for the shuttle is sufficient to justify paying the increased costs recently announced by NASA; but whether it is also prepared to support new science programmes remains uncertain.

Senator Adlai Stevenson Jr., chairman of the Senate subcommittee on science and space, warned at a hearing of the subcommittee last week that the shuttle problems could require an extra \$500 million a year from the US taxpayer to maintain a civil space programme. "And that may prove to be too high a price to pay," he said.

The most direct manifestation of the shuttle's problems have been the delays caused by problems in engine testing and, more recently, in fixing the heat-protective tiles to the outer skin of the orbiter. NASA officials now admit that the shuttle is unlikely to be airborne before next spring, perhaps even as late as the summer: and Dr Robert Frosch, NASA administrator, told the Senate subcommittee that total cost overruns for the design, development, test and evaluation part of the programme were now expected to be about 15%—or roughly \$1 billion in current dollar values—compared to a 7% estimate last year.

The engine problems, however, have also had effects on the programme. For example the reduced performance estimates for early shuttle launches have led to erosion of engineering and weight margins in space science payload design. One programme to have been affected by this is Project Galileo, a twin probe and orbiter due to be launched from the shuttle in 1982, the former to provide the first direct samplings of the planet's atmosphere,

the latter to transmit data from an elliptical orbit through all regions around the planet, including at least 11 near encounters with its surface.

The complexities of the manoeuvring required of the two vehicles on reaching the planet have resulted in a slightly heavier design than first anticipated. Initial fears that this might require discarding some of the scientific experiments have now been circumvented; but the extra weight together with the reduced performance expected of the early space shuttle flights, have required significant design changes in the so-called Interim Upper Stage (IUS) which will launch the Galileo vehicles from the shuttle.

At present, assuming that the IUS development goes as planned, performance margins appear to be acceptable. "We seem to have got the weight problem under control, and the project itself looks in pretty good shape at the moment," project scientist Dr Torrence V. Johnson of the Jet Propulsion Laboratory in California said last week.

Any further problems, however, either in the IUS development or in the shuttle test schedule would be a serious matter. At present Galileo, which will be among the first operational shuttle launches, is planned for early 1982, to reach Jupiter in the middle of 1985. On this schedule the spacecraft would fly close to Mars,

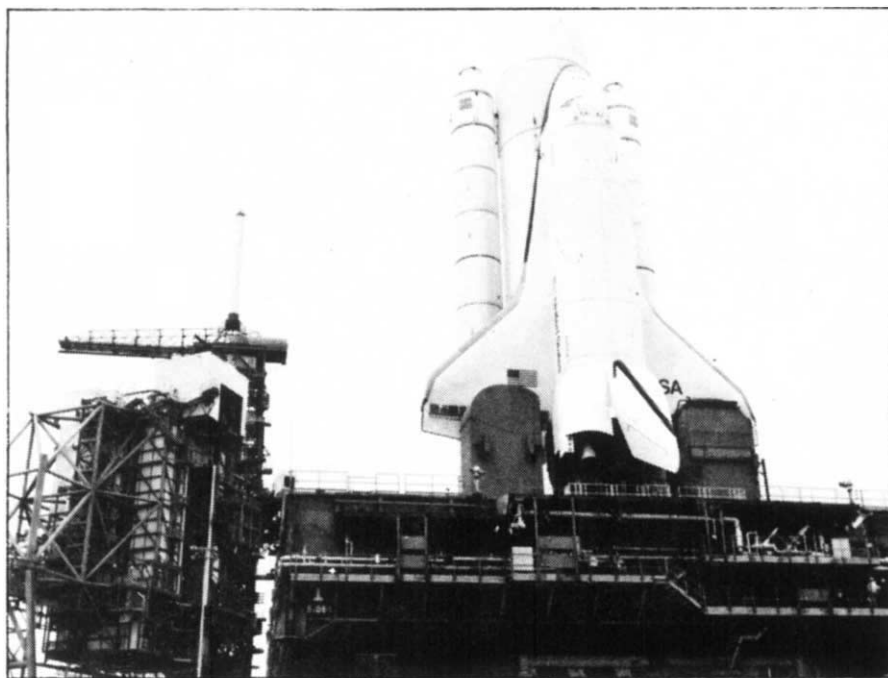
using that planet's gravitational pull to help it on its way.

A major delay in the shuttle programme, however, pushing the Galileo launch into 1983, would mean that the gravitational pull of Mars could not be used. Extra fuel would then be required to get the spacecraft to Jupiter, and some of the scientific projects would be dropped to compensate for the extra weight.

Various contingency plans are already being studied by NASA. One possibility might be to split the mission into two halves, sending the orbiter and the probe separately, perhaps a year apart; but this would involve considerable extra cost, and a major redesign.

Another project similarly affected by weight considerations is the international solar polar mission (ISPM). This is a joint project between NASA and the European Space Agency, due for launch at the beginning of 1983, which will involve sending two spacecraft simultaneously past Jupiter and into opposing orbits over the poles of the sun. Some of the potential weight problems for this mission have been compensated for by changing the trajectory of the launch from the shuttle. But the margins are now very tight (technically the ISPM is now 70 pounds—or 5% overweight, compared to 146 pounds for Galileo); and any problems with the IUS could have severe consequences.

As a contingency plan, NASA has



Space shuttle: setbacks are keeping other projects firmly on the ground

suggested that the two vehicles might be launched separately, possibly up to a year apart. But ESA scientists are not very happy with this suggestion, pointing out that a simultaneous launch is one of the attractions of the mission.

Weight problems are also causing concern to planners for the first Spacelab flight. This is still scheduled to take place in August 1981, a date that ESA officials are confident can be met even with considerable slippage in the shuttle test programme.

However as things stand at present, the shuttle may not be ready to provide enough thrust by this stage to launch the full Spacelab payload. The current shortage is about 1,300 pounds out of a total weight of about 65,000 pounds; and it could mean that some of the experiments already lined up for the first Spacelab flight may have to be dropped. (One encouraging piece of news to ESA officials is that the US Air Force appears to be planning to use Spacelab for some of its own manned space-flight missions.)

The space shuttle problems are already affecting other areas of NASA's activity. Scientists involved in the Voyager mission, for example, feel that if it had not been for these problems, they might have had access to some extra money this summer for additional experiments.

The greatest uncertainty at present, however, is over what the political effects of the shuttle's problems will be on support for future projects. NASA, with the support of both the Office of Science and Technology Policy and the Office of Management and Budget, is adamant that none of the existing programmes should be cut to provide the additional funds needed by the shuttle.

Dr Frosch told the Senate subcommittee that cancelling the Galileo project—at one time discussed as a distinct possibility—would only save \$83 million in 1980, considerably less than the extra money needed by the shuttle; nor would much be gained by taking money from the space telescope. "We do not consider either of these reasonable ways of getting extra funds," he said.

But the concern is that Congress, having provided NASA with the extra shuttle costs largely on the grounds that the success of the programme is required for US Defense plans—in particular for launching surveillance satellites, will be reluctant to complement this with support for new space science missions.

The situation is already bleak following the decision of the Carter administration not to request any new starts for the fiscal year 1980 partly to ensure adequate support for space

shuttle. (Among projects which NASA had submitted to OMB were plans for a shuttle-launched gamma-ray observatory, an orbiting imaging radar to Venus, and a mission to the comets Halley and Enke.)

"The real question is going to be what the 1981 budget contains. A second year with no new starts could really put the science programmes in jeopardy; if the nation wants a vigorous space science programme, then you have to start something new," Professor A. Cameron of Harvard College Observatory, chairman of the National Academy of Science's space science board, said last week.

Of course, if there are no new starts in the 1981 budget request the fault will not lie entirely with space shuttle; equally important may well turn out to be President Carter's desire to greet an election year with an attempt at a balanced budget.

But the additional funding which NASA is currently seeking from Congress for 1980 (\$225 million above its original request), has not improved the situation. "We are still hopeful that our project will be funded, but we are not as optimistic as we were a month ago," says one of the scientists involved with the gamma-ray telescope, currently top of the NASA's space astronomy and astrophysics board's priority list for future projects.

The dangers in the current situation have not been lost to Senator Stevenson who, with fellow committee member and ex-astronaut Harrison (Jack) Schmitt, has been urging a more aggressive space programme on the administration for some time.

"The problem with the shuttle programme now seems to have been a serious underfunding in the early stages, something of which up to six weeks ago the Congress had been unaware. And now it has all come home to roost with a 15% cost over-run," Mr Stevenson said at the subcommittee hearing last week.

"The biggest cost may effectively be the end of the civilian space programme, since the real risk is that the price of the shuttle's difficulties will be taken out of the civilian side of our space efforts. And with the country in its present mood, this could well happen."

Not all observers are quite so pessimistic. But the problems raised for future space science funding are already being taken seriously, not only by NASA but also by both OSTP and OMB. And they will provide an unwelcome house-warming present to greet NASA's new associate administrator for space science, Dr Thomas Mutch, when he takes up his appointment next month. □

California plans to build world's largest telescope

THE University of California is in the midst of preparing preliminary designs for a 10-metre diameter optical telescope that planners hope will be operating in the mid-1980s. Two designs are being considered, each a radical departure from conventional optical telescopes, which would be prohibitively expensive to scale up to this size.

The proposed telescope would be by far the largest in the world, with a mirror twice the diameter of the Hale telescope on Mount Palomar. (Nominally the largest telescope in the world is now the 6m at Zelenchukskaya Astrophysical Observatory in the Soviet Union, but western astronomers generally comment that it has not yet been proved reliable.)

Jerry Nelson, of Lawrence Berkeley Laboratory and chairman of the technical design committee for the proposed telescope told a meeting of Californian science writers last month that it would be designed to produce images smaller than 0.5 arc seconds and that, for good infrared observations, it would have to be built at a site with low atmospheric moisture.

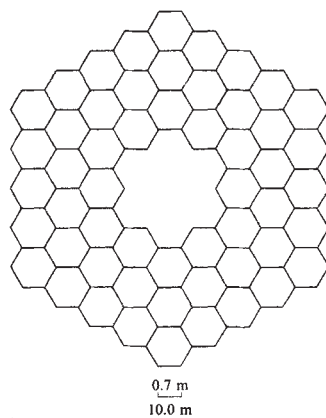
Historically, the costs of optical telescopes with thick, monolithic glass mirrors have risen roughly as the cube of the diameter. This means that a conventionally designed 10-metre instrument would cost around \$200 million—a figure Nelson says is "out of the question". On the other hand, the design committee has already rejected such alternatives as the use of independently focused multiple mirrors, like the recently opened telescope at Mount Hopkins (*Nature*, 3 May, page 9). Such multiple mirror devices may provide an economical alternative for studying point sources, but the new University of California telescope must have angular field wide enough to encompass an extended image like a galaxy with associated star clusters.

Nelson's design committee has thus narrowed the acceptable options down to two: a single mirror much thinner than those used previously, or a segmented mirror with focusing characteristics approximating those of the monolithic design. In either case the anticipated principal focal length would be between 15 and 20 metres, which would facilitate a relatively small containment structure. Traditionally half the expense of building a new observatory has gone into the telescope and half into the surrounding structure, so every effort is being taken to minimise the overall size.

The instrument will also be made more compact by the use of an altazimuth mount. Most large telescopes have been mounted with their axis of rotation parallel to that of the Earth, so that only one sweeping movement is required to track an object across the sky, but this so-called equatorial mount is relatively bulky. The advent of computer control has now removed the obstacles that once prevented use of the smaller but more complex altazimuth mount, which requires co-ordinated rotation about both the vertical and horizontal axes.

For either the thin continuous mirror or the segmented mirror, the question of how to evenly distribute the weight of the glass raises a host of unsolved problems. In either case 'active optics' will be required—that is, the stress on various support points must change continuously to prevent deformation of the glass when the mirror is tilted. The thin monolithic mirror would require some 100 to 200 support points, the majority of which must be active. On the other design, each segment of the composite mirror would require 12 sensors and 3 actuators to change its tilt and focus. On this point, at least, Nelson seems confident. "We think control is no problem," he says. A simple actuator device for segments of the composite mirror has already been tested and shown to be many times more precise than required, and the overall control scheme has been successfully modelled on a computer.

Preparing the glass, however, is another matter. A leading glass company has agreed to prepare the monolithic blank if that design is finally chosen, but many problems would remain in caring for the world's largest piece of fragile glassware. Among other things, the whole mirror, planned to be only 15 centimetres thick,



Possible segmented design for 10m mirror

would have to be enclosed in a vacuum tank each year for re-aluminizing. The segmented mirror also under consideration would require nine basic segment shapes to produce an overall parabolic curvature. The polishing might be done by computer control on segments already prepared for mounting, or possibly by a cheaper method—which Nelson calls "extremely promising"—in which the pre-stressed segments are polished spherically then released to snap back into the shape of an off-axis parabola. The ten centimetre thick segments would be hexagonal, 0.7 metres on a side, and mounted in three rings around a larger central mirror. The completed composite mirror would contain 54 segments.

The final cost should be in the range of \$20-50 million, of which the university hopes to raise at least half initially and then try to obtain matching federal funds. If construction can begin by 1981 the new telescope might be completed by 1986.

To avoid as much water vapour as possible, three relatively arid mountains have been proposed as possible sites. Mauna Kea, a 14,000 foot volcano in

Hawaii, would be the most accessible since the roads already lead to existing observatories on its peak. White Mountain in California, also 14,000 feet high, would be closer to the university's campuses, but the summit is almost unreachable by land vehicle. Junipero Serra Mountain is close to the UC Santa Cruz campus, which has a particularly active astronomy department, but with a height of only 6,000 feet it is by far the most moist of the three locations. Wherever the new observatory is built, the planners have already resigned themselves to building a structure sturdy enough to allow observations in the midst of 50 mph winds.

Nelson and others involved in the project sound optimistic about being able to build their proposed observatory, despite its high cost and necessarily radical design. They point out four broad classes of observation that would be facilitated by the proposed instrument:

- With direct optical imaging, a 10-metre telescope would permit better study of distant galaxies and a further attack on the problem of whether the universe will expand for ever or eventually collapse.
- Direct infrared observations would be used to study protostars and the structure of the galactic centre—perhaps resolving the question of whether it contains a massive black hole.
- Optical spectroscopy using the new instrument would allow astronomers to study globular clusters around other galaxies and investigate why quasars have absorption lines.
- Finally, infrared spectroscopy could be used to determine the abundances of various molecules in stars, and to study the emission spectra of quasars.

John Douglas

Sussex students step up science exam protest

PROTEST escalated last week at the campus of Sussex University against a compulsory first-year preliminary science examination. Two students, Richard Flint, president of the student union and Shaun Fensom, an engineering student, were "excluded permanently" from campus by vice-Chancellor Sir Denys Wilkinson for their "leadership" of a student-union-mandated disruption of a resitting of the contested examination.

An emergency Senate meeting has been called for 13 June when the Vice-Chancellor will explain his action to the faculty.

Half the university's 4,000 students attended an emergency meeting on 5 June. They rejected a call for an

immediate occupation of university buildings, but voted to canvass university and staff to protest against the expulsion and seek support from campus trade unions. They also decided on a series of one-day strikes the first of which was held last Friday with mass picketing at university entrances. The vice-chancellor was turned away. Students estimated that attendance was reduced to 10%-20% of normal but administration officials claimed that the action had not achieved its aim of stopping all classes.

The protest is taking a form not familiar since the 1960s. The original campaign began two years ago when the student union voted for direct student action to end a compulsory

first-year examination in science subjects. The science division is the only one at Sussex with such a requirement. Arguing that the exam distorted teaching practices and served no useful assessment purpose, the union called for a boycott. In the first year the boycott was 65% effective.

In the meantime a nine-member faculty working party was set up to consider the students' demand. It rejected the students' proposal that the exam be made optional, citing a need for "rigorous training in discipline", "an ability to marshal facts under pressure and to meet deadlines" and the "necessity for an objective means to monitor progress" as reasons for its decision. The working party specifically

rejected the students' request for tutorial assessment.

Students maintained their position, arguing that skill at examinations is unrelated to the skills it takes to be a competent scientist and that gearing the course towards quantitative assessment actually reduced learning. In January, a second boycott was called, with the addition that any

resitting of the exam that included boycotters would be disrupted.

The second boycott was far less successful with only 10% of the students failing to attend. Nevertheless the university rescheduled the exam to include students who had participated in the boycott. The exam was disrupted by 25 students who occupied the room. A second resitting was re-

scheduled, this time with faculty, police and photographers in attendance, but the students banged on tin cans and drums outside the examination room.

On 1 June, the two students were summoned to the vice-chancellor's office and expelled for their part in the disturbances.

Joe Schwartz

Soviets get their fun from learning about science

SOVIET society is in many ways surprisingly Victorian—a trend shown in the great popularity of the scientific lecture as a form of cultural entertainment. The organisation of such lectures is the task of the *Znanie* (Knowledge) Society, founded in 1947, which, in Soviet society, plays roughly the role of the British Association for the Advancement of Science, the Royal Institution, and the Worker's Educational Association in the United Kingdom.

Recently, two notable representatives of *Znanie*, Nobel Laureate Academician Basov and Dr Vladislavlev, toured the UK as part of the "Days of Soviet Science and Technology" organised in connection with the Soviet Exhibition at Earls Court. A delegation of 11 scientists and some 10 support personnel took part in the "Days". The host was the British Association—although, by some failure in communication, they were not aware until the delegation arrived of the great significance which Soviet science publicity puts on the organisation of such "Days". Nevertheless, one outcome of the visit was agreement to exchange lecturers between *Znanie* and the British Association—although it seems unlikely that the Soviet public will have the direct benefit of words of wisdom from the BA in the foreseeable future.

For *Znanie*, although lauded in official handbooks as being built on "democratic principles", is a hierarchical organisation. Although the society is supposed to organise lectures for any Soviet collective farm, factory or club that demands them, it would be impractical for top-flight lecturers to be constantly commuting, say, between Moscow and Vladivostok. "What we do", explained Vladislavlev "is to organise lectures for lecturers' in such remote places. Then the lecturers we have trained can give lectures at the local level".

If a top specialist is required, then the organisation requesting him must pay his fare. This suggests that the remote areas are somewhat dis-

advantaged—an idea which the visitors resolutely denied. As far as the Virgin Lands, and the far north of Siberia are concerned, they explained, there are special funds—and for the rest, there are no problems. Local lecturers are well trained to do their job in popularising science.

Lecturers, in fact, form the only 'members' of *Znanie* (apart from administration personnel). The main privilege of membership is to give frequent lectures (usually gratis) to whatever audience demands your services, and to attend frequent seminars on the latest developments in your field and the newest methods of lecturing. "The main problem", Vladislavlev explained, "is to make lectures appropriate for different audiences. That is why it is obligatory, when you ask for a lecturer, to describe the intended audience."

Not all lectures are simply a cultural diversion. *Znanie* is also responsible for the 'People's Universities', a fairly new feature of Soviet life. Courses are organised by plants and ministries to provide workers with the means of improving their professional qualifications.

Znanie itself is only concerned with the academic side—it is up to the factory commissioning its services to provide the facilities and to organise funding. Some 60% of *Znanie* funds, he added, come from non-state sources—factories, and trade unions. Of course, if a worker simply want to study, say, English Literature, he added, "then he pays for himself, just as if he was going to the cinema!"

To the Soviet mind, 'science' implies, as Academician Basov put it, "all advances in knowledge". Thus the 'flow-diagram' of *Znanie's* work includes 'Houses of atheism' in parallel with its planetaria. The most popular lectures organised by the society, according to Vladislavlev, are those on international affairs, followed by medicine.

Only the lecturers and organisers, not the audiences, are members of *Znanie*. If, say, some Ukrainian collec-

tive-farm wants to start a 'Bird-watchers' club', *Znanie* will provide lecturers and suggest reading matter, but the club cannot become an associate member of *Znanie*. Only high-level professional societies, such as the "Popov Radio-technical Society" has that privilege. (Some 40 such societies at present are 'collective members' of *Znanie*.)

Even so, with more than three million active members (lecturers and paid local organisers), there is no need to invite further people to join. "We have too many members already", Basov said. Some prominent personalities, of course, are virtually *ex officio* members. All Soviet cosmonauts, said Vladislavlev, are exploited enormously. "In fact, their bosses are always telling us to leave them in peace to get on with their job."

"There is an increasing demand for scientific knowledge and an increasing interest" said Vladislavlev. "The more one knows, the more one wants to know." As for the social impact of science, Academician Basov reiterated an earlier communication to the British Association that "science is not always good!" It would appear, however, that, in spite of a vast quantity of general science reporting (all major discoveries are reported in *Pravda*, Basov said), the Soviet public have not developed a "doomwatch outlook".

"It is *not* disasters that interest the public in science", said Vladislavlev, referring to Harrisburg. "The interest comes naturally from the rising cultural level."

And, as far as the Soviet citizen is concerned, everything is done to encourage that rise. "First we offer him a single lecture", said Vladislavlev. "Then, having caught his interest, a course of lectures. And finally, enrolment in a People's University. We are concerned in what you would call continuous adult education, and its development is our main task for the future!"

Vera Rich

news in brief

MRC staff committee prepares for budget fight: Responding to a leak in *The Observer* (3 June) predicting extensive cuts in UK science in this week's budget, the Medical Research Council Staff Side Committee initiated actions last week to try to prevent the cuts. Petitions and telegrams from every MRC unit have been sent to Neil Macfarlane, Minister of Science, protesting cutbacks in research funding. Eight of the nine trade unions represented on the staff committee supported a move for an immediate meeting with Macfarlane but the minister refused the request calling it "premature". Staff side initiatives to make joint representations to the government with MRC management were similarly rebuffed last week. In a letter, MRC management rejected the possibility of joint action opposing budget cuts saying that management and staff should act independently since the two have different views. The full staff side committee will meet 20 June to make its long term plans to prevent damage to medical research in the UK.

US scientists report successful laser fusion tests: Scientists at the University of Rochester in New York suggested last week that the highly successful results of a series of experiments on a laser fusion energy system might bring forward the date at which break-even in fusion energy—the point at which the energy produced by the system equals the energy inserted—can be achieved.

The experiments were carried out using a six-beam ZETA laser system, the first stage of a 24-beam system scheduled to begin operation later in the year. Using a laser power level of 1.65×10^{12} watts, the experiment generated more than one thousand million neutrons, at temperatures of 67 million degrees. "This is between five and ten times the amount that was expected," Dr Moshe Lubin, director of the university's Laser Energetics Laboratory, said last week.

When the full 24-beam system is in operation, it is hoped to be able to produce 30 to 40 terawatts ($\text{watts} \times 10^{12}$) of power. "The success of our first experiments means that it may be possible for us here at the Laser Energetics Laboratory to help the national programme move up the date that is projected for breakeven. That event is presently estimated to occur in the 1984 timeframe," said Dr Lubin.

Third World energy experts to confront the West: A symposium to be held in London 20–22 June will provide a forum for 30 spokespersons from the Third World to confront "the dominance being exercised by developed countries in science and technology, monetary expenditure, and planning and policy." The conference, titled Third World Energy Strategies and the Role of the Industrialised Countries is being organised by Ariane van Buren and Gerald Foley of the International Institute for Energy Development and is designed to give experts from the Third World a chance to address influential figures from western aid agencies and development banks in the areas of nuclear power, the role of multinational oil companies in development and the desirability of the alternate technologies currently in western vogue as being "appropriate" for the Third World. Papers will be pre-circulated and speakers will be restricted to five minute presentations thus allowing for up to one hour of debate. The papers and debates will be published. Sponsors of the meeting include the UK Ministry of Overseas Development, Atlantic Richfield, and the Beiger Institute of the Swedish Academy of Sciences.

UNCTAD bodes ill for UNCSTD: The increasing economic crisis facing the developed capitalist countries has produced a complete deadlock at the fifth UN Conference on Trade and Development. Four weeks of negotiations at Manila among 5,000 delegates from the world's rich and poor nations costing an estimated £100m have failed to reach agreement on any restructuring of the western world's economy. "The resolutions passed commit us to absolutely nothing" said one western delegate. Third World countries attempted to change current practices of the International Monetary Fund, to install a system of trade preferences and multilateral trade negotiations for developing countries and to break the industrialised nation's monopoly on shipping. Final resolutions condemned protectionism in trade against Third World products, provided for short and long term "action programmes" and expressed a need to progress towards an aid target of 0.7% of GNP. But several countries, including the UK, said they could make no commitment due to current cutbacks in public spending. The failure of the rich nations to yield to Third World demands for economic change reduces the possibility for effective action to be taken at the UN Conference on Science, Technology and Development to be held in Vienna in August.

India extends ban on wildlife export: The Union Ministry of Commerce has restricted wildlife exports to only 21 live species of birds, one live species of animal and products of six other animals, on the recommendation of a sub-committee of the Indian Board for Wildlife which examined the existing export policy from an ecological viewpoint. The only animal species that can now be exported is the striped squirrel; allowed bird species include myna (other than hill myna), parakeet, and two Himalayan species known for their colourful plumage—Sibia and Siva. Among the allowed animal products are the skins of common fox, hill fox, jackal and jungle cats, and peacock feathers and porcupine quills. The new policy follows an import-export order of January 1978 which banned most of the wildlife in demand abroad, and the recent revision will further hit the trade. The bird trade is worth 80 million rupees annually, the major portion being export: according to a recent report of the Royal Society for Protection of Birds, 4.5 million birds used to be exported from India every year. Export of munias, constituting 70% of the total bird trade, was banned last year, and the ban on the export of rhesus monkey (about 20,000 annually) and on the lucrative trade of pelts and furs of cat family and skins of lizards have drastically affected the animal trade. But there has been a concurrent increase in poaching and smuggling of wildlife and their products. Recently a global racket involving the

exchange of wrist watches and other luxury items for snake skins was uncovered in Bombay. So although conservationists and naturalists are pleased by the recent directive—except for jackal skin, an animal fast decreasing in number—they feel much more rigorous policing is required.

from Dilip M. Salwi,
New Delhi.



"Keep your part of the bargain.
Throw out the wrist watches!"

Commercial break for Britain's science?

THE word 'atmosphere' liberally peppers the speech of Neil Macfarlane. It is used constantly in attempts to describe just how important good morale is for the country's scientific community, and in particular, the newly-appointed Minister for Science believes there must be infinitely better links between "the research atmosphere" of the good universities and industry in Britain.

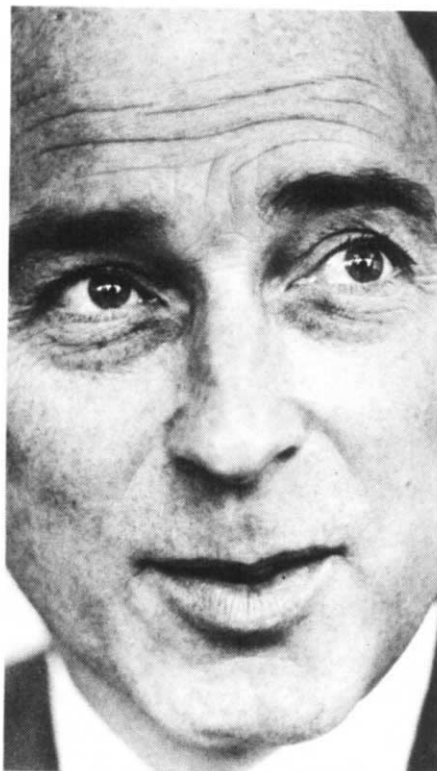
"We have lost a lot of ground over the past 20 years now", he states. "Too many industrial organisations have turned to their companies' research activities when looking for savings on their annual budget." To rectify matters, he maintains that we must encourage companies to set up centres close to universities—as the Americans do. Similarly our universities must be more commercially minded in looking for closer ties with major firms.

But 43-year-old Mr Macfarlane stresses that the real effort must come from industry, which in the past has shown little regard for scientific and engineering talent. "I think it is greatly to be regretted that the engineer and the scientist are generally relegated to a secondary level in the hierarchy of any major industrial organisation in the UK. We have been taken over by financiers and lawyers."

These are candid words from a junior minister who has no scientific training and whose background is solidly commercial. Educated at Bancroft's School, Woodford Green, he was commissioned in the Essex Regiment from 1955 to 1957. He left "after Suez, the Middle East and all that particular cauffle" for the oil industry, to eventually become a sales manager with BP; he was closely involved in the 1973 oil crisis, dealing with the distribution of petrol in Southern England.

In February 1974, he was elected Conservative MP for Sutton and Cheam and straightaway sat on the Select Committee for Science and Technology. His experience there opened up a completely new interest in science and Mr Macfarlane is quick to defend the committee's record for giving a good overview of science and technology and for its many short, sharp investigations: "I think the whole atmosphere which then developed because of the decline of our industry prompted many of us to pursue various technological issues with vigour and at all levels."

He is certainly no supporter of the proposal by the Select Committee on Procedure for the dissolution of the science and technology group. "I



Neil Macfarlane (above), Minister for Science in the new Tory government, wants closer ties between industry and research

Profile by **Robin McKie**

would be very apprehensive if it was entirely disbanded because I would envisage a tremendous amount of duplication of its work by other government departments." But defenders of the committee cannot look to him as an ally in their fight for its retention, because he maintains that, as a member of the government, he has no right to intervene in a purely House of Commons matter.

Even if the committee is disbanded, Mr Macfarlane will remain optimistic, believing that in general both parliament and government serve science well. "I just wish that we could have a little more motivation of other Members of Parliament towards the importance of science. There is not nearly the interest that there should be." He believes the previous Labour administration did not help by giving all responsibility for science to the Secretary of State for Education and Science, Mrs Shirley Williams, without delegating to a junior minister. Education became such a political football that it took up most of her time, at the expense of science.

"However, I am not going to

criticise her all the way because I think in many respects she did show a tremendous interest and grasp, and doubtless she would have liked to have spent more time on science. But now there is someone who can spare that time and has access straight through to the Secretary of State who is totally involved himself. We have got to ensure that the science fraternity knows it is not neglected by this administration."

This desire is reflected in the breadth of Mr Macfarlane's role, which covers science, research, technology, engineering and preparing for the advent of the microprocessor revolution. Although there have been junior ministers for science before, they have not been faced with the same sweep of responsibilities.

One major problem is staff stagnation at university science departments, which is freezing promotions, preventing staff exchanges between centres and hindering the cross-fertilisation of ideas. Mr Macfarlane refuses to be pessimistic, however. "I don't entirely take a gloomy view of staff stagnation. It was probably bad eight or 10 years ago, maybe even four or five but I believe there has been a great awakening over this problem in recent years. Things are beginning to move and there is this exchange and cross-fertilisation taking place. We have just got to accelerate it somehow."

In general, he foresees "a period of stabilisation" for universities, with the emphasis being placed on training the 16-19 age group to provide skilled technician manpower for our industrial re-birth. The research councils can expect to continue with the level of funding which has occurred in the past couple of years.

On the 'horny question of lay and union representation on the laboratory safety committees the Dangerous Pathogens Advisory Group and the Genetic Manipulation Advisory Group, Mr Macfarlane takes the same steadfast line as his Secretary of State, Mr Mark Carlisle, in opposing such moves. "These two bodies are so highly specialised that I don't believe there is a role for anybody else on them at this embryonic stage of their activities."

The real problem is using science properly in our industrial growth. "The main thing at the moment is that the atmosphere is right in the science departments—whereas five to 10 years ago it was very different. What I want to do is make Britain proud of its science, and that is going to be quite easy because our science is second to none. It is the very best." □

How French 'postdocs' are left out in the cold

UNIVERSITIES have stopped growing in France, as in other western countries, and young scientists who want to make careers in research are feeling the pinch as badly as anywhere else. One paradox of this situation is that, at a time when postdoctoral fellowships are becoming scarce, researchers are increasingly reluctant to accept temporary posts, and less willing to change jobs merely to continue in research. While the problem is international, its French version has some special idiosyncracies.

Research in France is not as university-based as in some other western countries. A major role is played by the Centre National de la Recherche Scientifique (CNRS), which despite its name is not a place but an institution, financing a large number of research laboratories throughout France. Some of these are located on university campuses and some in the Grandes Ecoles, while others are free-standing. A fairly close analogy would be with the MRC research units in Britain.

Purely academic research is financed partly through the Ministry of Education and partly through the Delegation Générale à la Recherche Scientifique et Technologique (DGRST), a body which fulfils some of the functions of the UK Science Research Council. The DGRST provides a limited amount of finance for hardware, a fairly generous supply of research studentships, but only a very sparse ration of postdoctoral fellowships, a function which largely rests with CNRS. But CNRS does not provide fellowships for academic departments, only for its own laboratories and those 'associated' with it (which together accounted for some 7,600 research workers and 13,400 technicians in 1977). Since it is becoming steadily more difficult for university professors to raise funds for postdoctoral workers, as described later, there is a risk of prolonged contraction of academic research as compared with the CNRS-supported laboratories. DGRST is not now, it seems, in a position extensively to support either university research or that in the Grandes Ecoles, in the way SRC does in Britain by means of research grants (which often include funds for graduate research assistants), research fellowships, advanced fellowships and the like.

In France, all academics with teaching posts, and also employees of bodies like CNRS, are *ipso facto* civil servants (*fonctionnaires*) and as such have permanent tenure. They do not need therefore, and indeed are not entitled, to take part in the state's unemployment insurance scheme, ASSEDIC.

ASSEDIC was designed to cater for workers in private industry or commerce; it has been the subject of much negotiation between government and unions, and its terms were most recently improved last March. It is now a thoroughly comprehensive scheme—with one yawning omission: it does not cover temporary research workers ('postdocs') in universities, Ecoles, CNRS laboratories or similar establishments. French law is quite firm about this: postdocs are neither *fonctionnaires* nor are they employees in private industry, and so they slip neatly between the available stools.

French postdocs were not particularly concerned about this state of affairs until 1971 or 1972, when scientific unemployment first became noticeable. Pressure then began to be exerted on behalf of the postdocs by their

6

An unfortunate quirk of French law, probably dating back to the Napoleonic concept of university, provides no proper legal status . . .

9

trade unions, such as the Syndicat National de Chercheurs Scientifiques, for temporary research workers to be eligible for unemployment insurance. This was not conceded by the government, but a number of temporary postdocs were converted into a hybrid species, *fonctionnaires hors statut* (literally, civil servants outside the law), and as such attained tenure. DGRST found itself obliged to meet the cost of these new *fonctionnaires* and this limited its resources for maintaining posts for those who were still temporary, let alone creating new posts. For this reason, among others, scientific unemployment grew further. The unions became very restive and successfully prevailed upon DGRST to terminate its long established *contrats de recherche* (broadly similar to SRC research grants) which often incorporated money to pay for postdocs. The main aim was to force government agencies to make all research posts permanent, but the government has done the opposite. It has reduced the number of temporary posts without increasing the number of permanent positions.

Before this happened, industrial concerns had been in the habit of supporting contract research in universities, and financed a number of academic postdocs; but anxiety among young scientists facing the possibility of eventual unemployment without

state unemployment relief as well as union hostility, caused the withdrawal of industrial contracts, and these do not now appear to be a significant source of funds.

An unfortunate quirk of French law, probably dating back to the Napoleonic concept of the university, which provides no proper legal status and thus no insurance for those who are neither civil servants, self-employed nor employees of firms in the private sector, thus has seriously added to the plight of untenured French research workers. Initiatives to set up an unemployment scheme independent of the state do not appear, up to the present, to have met with success.

DGRST, aided by the new science minister, has striven valiantly to fill the gap left by the atrophy of contract research. In addition to some 3,000 2-year studentships for predoctoral students, and some 60 fellowships for tenured academics to enable them to devote some years to fulltime research, DGRST has now begun to provide postdoctoral fellowships available for allocation in the principal research laboratories. But at the time of writing the number available through DGRST is only of the order of 100, which falls a long way short of meeting the need. The greatest hardship is among newly graduated PhD's.

An influential working group of academics, industrialists and administrators presided over by Professor Jacques Friedel has been taking steps to assemble funds for some university fellowships, which have begun to be available at Orsay, Marseille, Strasbourg, Poitiers and Lille; other fellowships have been established in the State laboratories at Saclay and Grenoble. The insurance deadlock has not been resolved, but at least the decay of contracts is beginning to be circumvented.

This situation is however exacerbated by another specifically French problem, the firmly established tradition of young scientists to remain in the department where they researched their theses. Several administrators expressed the view to the author that the difficulty in persuading young scientists to move to other laboratories or to neighbouring disciplines, is the greatest single obstacle, after the insurance laws, to fuller scientific employment; and from a purely scientific standpoint, immobility has the grave disadvantage that the inflow of new ideas and experience into long-established laboratories is prevented.

R. W. Cahn

The author is Professor of Material Sciences at Sussex University.

Snail-paced parasite that is marching through South America

Schistosomiasis, otherwise known as Bilharzia, is a debilitating disease causing anaemia, diarrhoea, abdominal pain and sometimes death. Its spread, initially through a particular species of freshwater snail, is now being accelerated by human carriers—Brazil's millions of migrant workers

Report by David Bousfield of Sussex University

UNTIL the 1920s schistosomiasis, a debilitating parasitic disease caused by *Schistosoma mansoni* which is transmitted by certain species of freshwater snail, was more or less restricted to the north-east of Brazil—an area which includes the states of Ceará, Rio Grande do Norte, Paraíba, Pernambuco, Alagoas and Sergipe. Then, as now, sugar cane cultivation along the coastal 'zona da mata' was a major source of employment. The sugar mills or 'engenhos' were invariably associated with river valleys providing alluvial fertilisation during the winter, water for irrigation, and ideal habitats for *Biomphalaria glabrata*, the main intermediate host of schistosomiasis in the area, which is capable of adapting to almost any kind of freshwater environment. Inland the climate becomes progressively drier and the corresponding vegetational zones, the 'agreste' and the 'sertão', supported simple farming and cattle ranching. In these semi-arid regions limited and intermittent sources of water meant concentration of population, pollution and disease and provided an environment for an intermediate host, *Biomphalaria straminea*, that was well adapted to temporary desiccation. Today, estimates for the levels of infection in Sergipe and Alagoas are above 30%, and even in the most developed state, Pernambuco, the level was almost 16% before 1975.

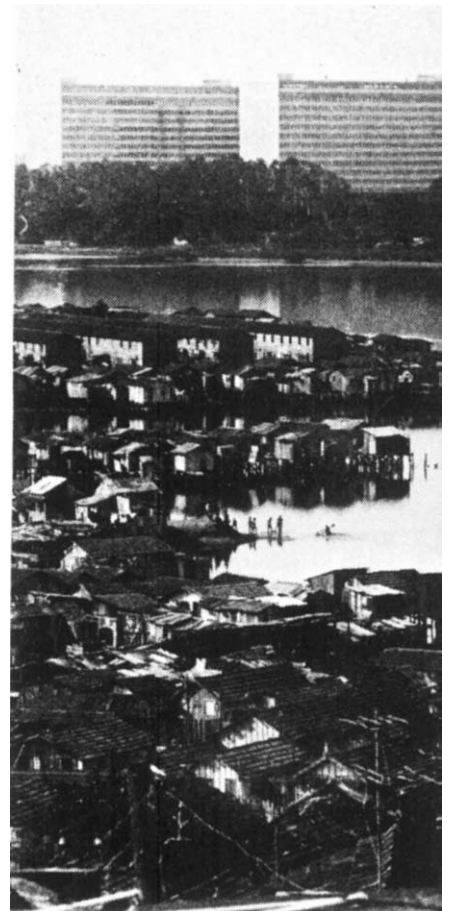
Unfortunately the disease did not remain in the north-east. Competition from the more progressive plantation owners in the south closed many 'engenhos' and forced others to adopt more mechanised, less labour intensive methods, whilst in the interior the severe droughts of 1958 and 1970, together with the 'strong-arm' tactics of the richer landowners forced many 'nordestinos' off their land and on to the road. These migrants travelled mainly to the south in search of work, taking their diseases with them. During the 1970 drought, for example, out of the total population of 30 million, 2 million people left the north-east. In 1937 schistosomiasis reached the state of Minas Gerais, and the states of Rio de Janeiro and São Paulo in 1950. In 1953

the disease reached Paraná bordering with Argentina and Paraguay, and it was recorded in Brasília in 1973. Since potential intermediate hosts were already to be found in all of these states it is presumed that human migration was responsible for the spread of the disease. By 1975 it was estimated that there were between 8 million and 18 million carriers and the disease cost \$150 million per year.

Human migration, however, is not the only reason for the rapid increase in size of the endemic area. Until about two decades ago only *Biomphalaria glabrata* and *B. straminea* were recognised as being the important intermediate hosts for schistosomiasis in South America. A third species *B. tenagophila*, found only in the south of the country (but whose distribution stretches westward and southward into neighbouring Bolivia, Paraguay, Argentina and Uruguay) was thought to have a very low susceptibility to infection by the miracidia of *S. mansoni*. Indeed in 1923 the Brazilian taxonomist Adolpho Lutz had proposed that the species be called *Planorbis immunis*. In 1956, however, Lygia Corrêa and others found a number of foci in the Paraíba valley which were maintained solely by *B. tenagophila*—in some areas snail infection rates were as high as 48%. Subsequent laboratory work conducted by Lobato Paraense and his co-workers have shown that *S. mansoni* is gradually adapting to its new snail host and that in the case of the strain from the Paraíba valley this process is particularly well advanced.

Will the strain spread to Amazonia?

In a recent paper by Paraense and Corrêa they remark that "the susceptibility of most *B. tenagophila* populations to the S. J. [Paraíba] strain points to a fact of practical significance—the possibility of expansion of that strain to a wide South American area, most of which is still free of schistosomiasis". They also report that the strain is gradually spreading, and that this rate will increase as stronger migration incentives and better economic opportunities are developed in the region.



Brazil's waterside shanty towns: a focus for transmission of disease

In addition to this movement south, spread of schistosomiasis has been recorded, albeit on a much smaller scale, in the northern states of Maranhão and Pará. The intermediate host here is *B. Straminea* but a new potential host *B. amazonica* has recently been discovered to the west with a distribution stretching from Porto Velho, along the Rio Solimões to Manaus. The possibility of spread to Amazonia is generally acknowledged amongst public health workers to be small. This judgement is based on the low population densities and high flow volumes involved, and the unsuitability of the water chemistry. However, the development of the Transamazonian Highway, built specifically to act as a 'safety valve' for the drought-plagued northeast, must challenge some of these assumptions. The original project, announced in 1970, called for settlement of 100,000 families along the pioneer highway by 1976—although, in fact, by September 1977 only 5,333 had actually been settled. Small towns were constructed every 10 km (agrovilas) with larger administrative and manufacturing centres built at 100 and 400 km intervals. Research conducted by Dr Nigel Smith (INPA, Manaus), however, has shown that as of September 1974 only 4 out of 26 agrovilas were equipped with piped water and

only 10 out of 26 possessed privies. Gastrointestinal problems due to the use of water from streams and ponds contaminated by human and animal faeces are already one of the major reasons for hospital admission. SUCAM, the government organisation responsible for disease control, maintains surveillance at Marabá, Altamira and in Rondônia, but to date no transmission has been reported, although many of the settlers are infected. However, at SUCAM's headquarters in Brasília, Dr Solon Camargo pointed out to me that agricultural activity and other attempts to modify forest ecology might produce favourable snail habitats in the vicinity of the highway. According to Dr Smith the 'slash and burn' techniques of the settlers and the application of lime to the soil has produced a large modification in the chemistry of stream and pond water close to the road and at a number of sites, notably Altamira *B. straminea* populations already thrive.

Until 1975 attempts to control Brazilian schistosomiasis took the form of small pilot projects usually aimed at reduction or eradication of snail populations using molluscicides. Some attempts to reduce the incidence of the disease by chemotherapy had also been made but neither strategy had been particularly useful in controlling transmission—small residual populations of infected snails could still maintain high rates of transmission and treated patients soon became reinfected, although it might be several years before disease reached its more advanced forms again. Indeed by this time many Brazilians working in schistosomiasis control felt that attempts to reduce the disease should be linked to community health programmes, and in particular to the provision of piped water, toilets and communal laundries. In rural Brazil

75% of households were without any form of toilet and 71% without protected water supplies in 1975. Dr F. S. Barbosa, a specialist in snail control wrote in 1974 that "attempts to control schistosomiasis, in a large scale, by using the conventional single control methods are futile and, more than this, may become detrimental to the developing communities."

Better sanitation key to prevention

When the first national plan to control schistosomiasis (PECE) was announced in 1975 at the beginning of the term of office of President Geisel it was greeted by considerable criticism from the scientific community, for the Minister of Health, Dr Paulo de Almeida Machado proposed the eliminate, or at least control, the disease by mass chemotherapy—the drug oxamniquine (Mansil) being used most extensively. This choice of strategy was based on the success of a small pilot scheme in Paraná and the success of an earlier mass injection scheme in which 80 million people were vaccinated against meningitis in 10 months. It is probable too that the need for positive results before the next election in 1978 played a part in the design of the programme. PECE began with two closely monitored experimental projects—one at the town of Touro (population 2,250) in the state of Rio Grande de Norte, and the other at Santo Antonio das Trempes (pop. 571) in Pernambuco. At Touro almost all the population was treated 'en masse' with oxamniquine during December 1975. Infection levels dropped from 56% to 5% but rose to 19% after one year. Indeed, for children under 15 years, the highest risk group, more than 50% became re-infected during this period. On the other hand, at Santo

Antonio das Trempes where chemotherapy was preceded by mollusciciding and installation of piped water, toilets, showers and communal laundries, the infection rate, which originally had been 50.4%, was still only 2% twenty months later. PECE has since been extended to cover seven north-eastern states, yet still relies heavily on drug therapy. Oxamniquine will be given to 12 million people (already between 1 and 2 million have been treated) whereas provision of sanitation is planned for only 2.6 million.

Additional programmes designed to control and stem the spread of schistosomiasis in the south are being undertaken by the São Paulo Secretariat of Health and in the neighbouring state of Paraná by SUCAM. São Paulo's problems began with its rapid industrial and agricultural growth with the concomitant need for manpower, irrigation systems and power stations. For example the state is now the country's biggest producer of sugar cane, a crop whose importance to the national economy will increase rapidly as the government's programme to substitute ethanol, produced from sugar, for gasoline gets under way. The need to reduce energy imports has also made hydroelectric power an increasingly attractive alternative and new schemes, such as the one at Itaipú, now have catchment areas extending into Argentina and Paraguay, making international transfer of schistosome strains a possibility. During the period 1967-70, the state received 113,054 migrants, mostly from the north-east and Minas Gerais. Routine coprological examination showed that an average 23% had schistosomiasis. By 1974 it was estimated that a quarter of a million infected migrants had entered the state. Attempts have been made to screen and treat infected migrants as they arrive. However, in order to prevent the spread of schistosomiasis such a scheme required that all migrants register with the local authorities, and that diagnosis and treatment are 100% effective. To date none of these standards has been achieved and the number of active foci within the state increases.

Whatever else happens in the future, it is clear that Brazil's need for cheap migratory workforces will continue to increase, and it is to be hoped that the new Minister of Health, Dr Mario Augusto de Castro Lima, with an extended 6 year term of office in front of him, will realise that chemotherapy is no match for the powerful social and economic forces which have made schistosomiasis a major problem in Brazil's agro-industrial age. □



Migrant sugar workers: attempts to screen them medically have failed

David Bousfield visited Brazil on a Nature Writing Fellowship

Which researcher will get the grant?

Jonathan R. Cole and Stephen Cole describe how disagreements among peer reviewers can explain the unexpectedly low correlation between a scientist's prior professional performance and his or her chances of being awarded a research grant.

In recent years the system used by the United States government to distribute basic research funds for scientists has been under increasing attack both from members of the scientific community and from the American Congress. The most important criticism that has been raised is that the current system based upon peer review leads to inequitable decisions favouring eminent scientists and putting non-eminent scientists at a serious disadvantage in the competition for increasingly scarce funds.

For the past three years we have been conducting a detailed study of how the National Science Foundation decides whether or not to approve applications for research funds (Stephen Cole, Leonard C. Rubin, Jonathan R. Cole, *Peer Review in the National Science Foundation* (Washington: National Academy of Sciences, 1978)). We have done this work as consultants to the Committee on Science and Public Policy (COSPUP) of the National Academy of Sciences. Our study is based upon extensive qualitative interviews with people involved in all aspects of peer review, particularly NSF programme directors; an analysis of the actual reviews given to 250 research proposals; and a quantitative analysis of 1,200 applicants to 10 different NSF programmes in fiscal year 1975. The purpose of the quantitative study was to identify those characteristics of scientists and their 'track records' that were correlated with high or low peer ratings and with the final funding decision.

'Old boys club' theory disproved

Although the NSF peer review system had been criticised as being an 'old boys' club', our research did not support the contention that the methods used by the NSF in deciding who should receive funding were inequitable. We found that eminent reviewers did not favour applications from their eminent colleagues. Furthermore we found weak correlations between measures of the scientific performance of the applicant and the peer review ratings they received. Finally we found that the peer review ratings did in fact almost completely determine whether or not a grant was funded or declined funding.



One important result of the first phase of our finding of the first phase of the research concerned the relationship of past 'track record' of the applicant to the peer review ratings they received. For example, we found that the number of papers published in the last 10 years and citations to those papers had only a slight influence on ratings given to a proposal by the peer reviewers. In two of the 10 programmes we studied the proportion of variance explained on ratings by citations were 0.16 and 0.14, and in the other programmes less than 0.10. Other productivity or output measures are even less strongly related to the ratings.

We also found that the prestige rank of the applicant's current academic department had only a small to moderate correlation with the ratings given to their proposals. The criticism that applicants from prestigious East and West coast universities have a significantly better chance of receiving grants than those from other universities turned out to be unjustified.

In this discussion our primary concern is *why* we did not find the expected relationship between measures of the scientist's past performance, his position in the stratification system of science, and the peer review ratings received on his proposal.

One reason why the characteristics of applicants showed such low correlation with ratings received was that there was an unexpectedly large amount of disagreement among reviewers of the same proposal. Some reviewers gave proposals high ratings and other reviewers gave the same proposal low ratings. If the same applicant gets high ratings from some reviewers and low ratings from others, the correlation between any characteristic of the applicant and the ratings must necessarily be low.

In order to estimate the level of

agreement among reviewers of particular proposals we performed a one-way analysis of variance for each of the 10 programmes. In this analysis, the dependent variable was the rating that a proposal received from a reviewer; the independent variable was simply the proposal number, which is of course a nominal variable. The results presented in the accompanying table (overleaf) show the proportion of the total sum of squares for ratings on all proposals that is attributable to disagreements among reviewers of the same proposal.

Lack of consensus is common to all subjects

Take, for example, the algebra programme at the NSF. If we examine all the variation that exists among raters of all proposals to the algebra programme, we find that 45% of the total variation is due to disagreements about the quality of individual proposals among the peer reviewers of the same proposal. The level of disagreement among reviewers of algebra proposals was by no means anomalous. We should also point out that contrary to our expectations there was no less consensus in the social sciences such as anthropology and economics than there was in the natural sciences. In fact, economics was the programme which showed the lowest proportion of variance (35%) explained by within-proposal variations.

These disagreements could be the product of several processes. First, there is substantially more actual cognitive disagreement among reviewers about the merits of proposed work by the applicant than is often believed to be the case. Second, the high amount of disagreement among reviewers could be in part a result of intersubjectivity, that is, two scientists might actually

Consensus among reviewers and influence of applicant characteristics on peer review ratings

	Lack of consensus among reviewers		Pearsonian correlation coefficients between				
	Percentage of total variance due to within-proposal variance	Log of citations to recent work and individual ratings	Log of citations to recent work and mean ratings	Rank of dept and individual ratings	Rank of dept and mean ratings	Combination of 9 characteristics and individual ratings*	Combination of 9 characteristics and mean ratings*
Algebra	45	.06	.11	.07	.10	.17	.31
Anthropology	43	.00	.01	.00	.00	.04	.07
Biochemistry	49	.16	.27	.07	.12	.20	.40
Chemical dynamics	45	.14	.25	.02	.04	.16	.29
Ecology	54	.01	.02	.02	.04	.06	.13
Economics	35	.08	.08	.13	.23	.21	.32
Fluid mechanics	43	.03	.02	.10	.10	.17	.29
Geophysics	57	.07	.14	.03	.07	.09	.21
Meteorology	63	.08	.16	.05	.11	.14	.37
Solid-state physics	55	.08	.16	.08	.16	.17	.38

* The nine characteristics include: professional age, rank of Ph.D. department, rank of current department, academic rank, log of papers published in last 10 years, log of citations to work published in last 10 years, log of citations to papers more than 10 years old, whether institution is a Ph.D.-granting one or other, number of years funded by NSF in last five years.

have the same substantive appraisal of a proposal but may use different subjective scales of evaluation. Consequently, a similar evaluation can result in different actual ratings. Third, the disagreements could reflect some personal bias that the reviewers have against a particular principal investigator who submits a proposal. We could catalogue other potential sources of variations in the ratings, but the important, albeit tentative conclusion to underline is that the amount of variation in peer review ratings that could possibly be explained by characteristics of scientists and their track records is distinctly limited by this disagreement among reviewers over the same programme.

Taking published papers into account

In order to eliminate the effects of this lack of consensus on the correlation between applicants' characteristics and ratings, we used as the dependent variable the mean rating received. As independent variables we used the log of citations to papers published during the last ten years prior to application, rank of the applicant's academic department, and a combination of nine characteristics measuring the scientist's location in the stratification system of science. The results of these regression analyses appear in the table.

Given that about one half of the variance in ratings is the result of within-proposal disagreement, the correlation between the characteristics of the applicant and the mean rating is about double the correlation between characteristics of the applicant and the individual rating. Although the correlations between individual characteristics and mean ratings are higher, they are still not as high as one might expect, particularly in fields such as solid state physics, algebra, or chemical

dynamics.

There are three reasons why these correlations are not higher. The first we call *self-selection*. Not everyone who is eligible to apply for an NSF grant does so. The group of scientists that applies for NSF research money tends to be a self-selected group of the most creative and productive scientists in their particular fields. In order to apply for a grant, an applicant must write a proposal that he or she knows will be sent out for review to colleagues in the discipline. Scientists who do not have active, ongoing research programmes are less likely to write such a proposal and subject themselves to peer review than those who do have active research programmes. Data in which we compare the productivity of the applicants to the NSF with productivity of random samples of academic scientists indicate that NSF applicants are on the whole a more productive group than are the population of scientists in a given field. Because self-selection takes place, the group of applicants from whom the reviewers must select successful proposals tends to have attenuated variance on some of the independent variables.

One reason for the low correlation between the rank of an applicant's academic department and peer review ratings is the relatively wide distribution of academic talent among American universities. Although it is true that, on the average, highly prestigious departments have more talented and productive scientists, a significant proportion of talented scientists are not located at the most prestigious departments. Several independent studies have found that the correlation between citations to a scientist's work and the prestige rank of his department is 0.30 or less. This suggests that quite a few scientists who have produced work judged to be useful and of high quality are not in highly ranked departments.

When we relate the low correlation between the quality of an individual scientist's research output and the rank of his department to the concept of self-selection, we can understand better the low correlation between the rank of an applicant's department and peer review ratings. If every scientist in every department applied for a grant, there would probably be a considerably higher correlation between rank of department and rating. But we know that all scientists do not apply. Applying scientists from low-ranked departments are probably the most active researchers of those departments. Whereas six mathematicians from MIT may apply for NSF funds in a given year, perhaps only one mathematician at a lower-ranked department will apply. But this one individual may have a national reputation comparable to those of some of his colleagues at high-ranked departments.

Quality of science is main influence

Finally, the third and perhaps most important reason for the low correlation between characteristics of the applicants and peer review ratings is that all evidence collected thus far in the peer review study suggests that the ratings are strongly influenced by the reviewer's perception of the quality of the science contained in the proposal. If this is indeed true and the correlations between the ratings and the characteristics of applicants are low to moderate, this would necessarily mean that the correlation between characteristics of applicants and the perceived quality of their proposals would be low to moderate. We are currently analysing a set of data which will allow us to determine the extent to which this hypothesis is correct and the meaning of these data for peer review and the distribution of federal funds to research scientists. □

news and views

Temperatures in the clouds of Venus

from P. J. Houghton

ON board the Pioneer Venus orbiter which encountered the planet Venus on 4 December 1978 was a multi-channel infrared radiometer jointly built by the Jet Propulsion Laboratory, Pasadena, California and the Department of Atmospheric Physics at the University of Oxford. The Oxford contribution to this radiometer is the first British built experiment to leave the vicinity of the Earth and travel to one of the planets. Remote sounding observations with this radiometer have already provided the first information on the temperature structure in the upper Venus atmosphere (Taylor *et al.* *Science* **203**, 779; 1979). On page 613 of this issue of *Nature* Taylor and his colleagues at the Jet Propulsion Laboratory report some interesting new measurements of the detailed thermal structure of the clouds near the Venus pole which provide important clues to the nature of the circulation of the Venus atmosphere. Investigations from Earth-based telescopes and from spacecraft have demonstrated the very deep, rather uniform cloud cover on Venus, which completely obscures the surface from view. That the cloud droplets mainly consist of sulphuric acid was discovered as a result of the interpretation by Hansen, Arking and Young in 1973 of polarisation observations made at different phases of Venus by the French astronomers Coffeen and Gehrels in 1969. Although the size of the particles, as also deduced from these measurements, turned out to be a few microns only, the question remained as to how the upward motion necessary to maintain the clouds was generated, even though for such small particles the amount of upward motion need only be very small. A further question concerned the location of the compensating regions of downward motion. This week's report gives support to the theory put forward by Murray *et al.* (*J. geophys. Res.* **68**, 4813; 1963) and Suomi and Limaye (*Science* **201**, 1009; 1978) that intense regions of downward motion might be concentrated near the poles. In these regions clearing of the clouds might be expected, which would

not be visible from Earth because the polar regions are viewed so obliquely.

The Pioneer 12 orbiter, because it goes nearly over the Venus pole, has provided the first opportunity to observe the polar regions in detail. The infrared radiometer observations clearly show a localised region of much higher temperature than the rest of the cloud cover which has been interpreted as a local clearing of at least the upper half of the cloud layer. The measurements do not rule out a clearing to much deeper levels.

Observations from the infrared radiometer at altitudes of 60–90 km in the Venus atmosphere, well above the main cloud deck, have shown a temperature contrast of about 20 K between equator and pole, with the pole being warmer (Taylor *et al. op. cit.*). The obvious interpretation of these measurements in terms of circulation is of a Hadley type circulation with rising air over the equator, cooling as it rises, and sinking air around the polar regions, warming as it sinks. This is not unlike the thermally driven circulation which occurs in the mesosphere of the Earth's atmosphere where very cold temperatures occur at the mesopause (~85 km altitude) in summer contrasted with much warmer temperatures near the winter mesopause in the other hemisphere. Another important feature of the circulation at those levels above the main cloud deck is the very rapid zonal motion which occurs. Air moves around the planet at a velocity of about 100 ms^{-1} , one rotation of the planet taking about 4 days. This rapid motion ensures that there is little temperature contrast between the day and night side of the planet, as is confirmed by the Pioneer radiometer observations.

The solid surface of Venus rotates only very slowly—once in 243 days—so that the comparatively rapid rotation of higher parts of the Venus atmosphere is particularly intriguing. Schubert and Whitehead in 1969 (*Science* **153**, 71) put forward a possible theory of the mechanisms driving this circulation; they suggested it was in-

duced by a travelling thermal wave induced by the motion of the Sun relative to the atmosphere.

At altitudes above about 90 km in the Venus ionosphere, the thermal contrast between equator and pole largely disappears. Instead at these levels, as is shown again from infrared radiometer measurements, a temperature difference of about 20 K exists between the day and night side of the atmosphere, indicating a mean circulation which is primarily occurring between the warmer sunlit side and the cooler dark side.

It will be interesting to discover, from further work on the data from the infrared radiometer and from the other instruments on the Pioneer 12 orbiter, whether there is any evidence for large scale waves in the atmosphere above the Venus clouds of a similar or different nature to the planetary waves observed in the Earth's atmosphere. A big question which clearly arises is: how are the circulations in different parts of the Venus atmosphere linked? Is the upper atmosphere circulation, for instance, driven from below, or is it driven by heat sources and sinks near the cloud tops? A particularly interesting observation is that the cloud clearing reported in this issue is not actually at the pole but some 10° from it, suggesting that the circulation may have some interesting asymmetrical properties. Clearly a lot remains to be understood that will not only encourage those working on Venus data to extract from it information which will provide clues to some of the intriguing problems of the Venus circulation, but also provide the theoretical modellers with a challenge to fit it all together. As we begin to understand how the motions are organised in the Venus atmosphere, which is so different from our own, we shall in turn acquire deeper insight into what goes on in the atmosphere of the Earth. □

P. J. Houghton is Professor in the Department of Atmospheric Physics, University of Oxford.

Neoplasias and B-cell precursors

from Maxime Seligmann

THE earliest recognisable cells within the B-cell differentiation pathway are known as pre-B cells. They are found in the fetal liver of rabbits, mice and men and are characterised by the presence of intracytoplasmic IgM and the lack of surface receptors¹⁻³. In adults, pre-B cells are found almost exclusively in the bone marrow⁴. IgM can be detected by immunofluorescence techniques in the cytoplasm of most, if not all, large rapidly dividing pre-B cells and their small non-dividing pre-B cell progeny⁵, with an unusual perinuclear and reticular staining pattern. Pre-B cells are the progenitors of surface Ig positive B lymphocytes but, because they do not express functional immunoglobulin receptors on their surface, it can safely be assumed that antigens do not influence the diversity of immunoglobulins produced by pre-B cells. Various studies suggest that clonal diversification is accomplished by the pre-B cell stage. For example, individual pre-B cells show allelic exclusion in rabbits heterozygous for kappa chain allotypes⁶. In myeloma patients, antibodies specific for idiotypic determinants of the myeloma globulin stain bone marrow cells with pre-B cell characteristics⁷; this finding provides evidence that pre-B cells belong to the neoplastic clone and suggests that V-C joining has occurred at this stage.

Several recent studies have shown that some murine and human neoplasias involve such early B-cell precursors which produce cytoplasmic immunoglobulin in the absence of membrane-bound immunoglobulin molecules. The Abelson murine leukaemia virus complex, which consists of a replication-defective virus and its helper Moloney leukaemia virus, causes thymus-independent lymphomas in Balb/c mice, and several characteristics suggested that these tumours were related to the B-cell lineage⁸. Two independent studies^{9,10} have recently provided evidence that several neoplastic cell lines resulting from *in vivo* or *in vitro* transformation by Abelson virus have features indicative of early maturation compartments in the B-lymphocyte lineage. On the other hand, two human lymphoid cell lines (Raji and T51) were also shown to express characteristics of early B-cell precursors¹¹. Finally the

leukaemic cells of some patients with acute lymphoblastic leukaemia also display features thought to be characteristic of pre-B cells. This pre-B cell phenotype is not uncommon because it was found in four of 22 patients in the initial report by Vogler *et al.*¹² and in 10 of 68 patients in the subsequent series of Brouet *et al.*¹³.

The results of immunofluorescence and immunoglobulin synthesis studies suggest that the maturation arrest in these various neoplastic cells may occur at slightly different steps in the early stages of the B-cell pathway. Early studies indicated that normal pre-B cells produce complete monomeric IgM molecules with both μ and light chain^{1,2}. In contrast, no light chain determinants could be detected by immunofluorescence in most human leukaemic pre-B cells¹³. This negative finding could be due to limitations of immunofluorescence technology because, in the human cultured cell lines mentioned above, light chains could not be detected by immunofluorescence but were found at the ultrastructural level by an immunoperoxidase method¹⁴. It should be noted, however, that the number of normal pre-B cells containing detectable light chain determinants is lower than the number of such cells with detectable μ chains¹⁴. This discrepancy and the findings in acute lymphoblastic leukaemia are consistent with the possibility that there is an asynchronous onset of heavy and light chain synthesis. The data reported by Siden *et al.*¹⁰ support this hypothesis because most Abelson-virus-induced cell lines synthesise a μ chain but no

detectable light chain. This μ -only phenotype was stable on subcloning. Cultures in which synthesis of both μ and κ chains occurred showed a rapid loss of κ chain synthesis. One line was studied in detail: the intracytoplasmic molecules represented glycosylated μ chain dimers which were not secreted. Light chain synthesis could not be induced by reducing the growth rate of the cells, a procedure that resulted in increased μ chain synthesis. In contrast to these findings, another Abelson-virus-derived line studied by the same authors produced only κ chains. It should be emphasised that the biosynthesis experiments showed that the cytoplasmic extract contained, in addition to the κ chain band, a slower band which comigrated with a precursor polypeptide. It is interesting in this respect that the intracytoplasmic IgM of the human lines was not located inside endoplasmic reticulum but at the level of free polyribosomes¹¹ and that these lines produce abnormally large Ig polypeptide chains (unpublished results of P. Guglielmi and J. L. Preud'homme). The two Abelson-virus-transformed cell lines studied by Boss *et al.*⁹ apparently synthesised monomeric IgM with both μ and light chains. However, most cells in both lines were presumably less mature than pre-B cells because they were negative for cytoplasmic IgM but could be induced to express IgM positivity by several compounds such as lipopolysaccharide, DMSO and butyric acid. The human lymphoid cell lines are presumably slightly more differentiated than *bona fide* pre-B cells because they bear Ia-like determinants, they secrete polymeric IgM molecules (instead of monomeric IgM in normal pre-B cells²) and they become able to produce membrane-bound immunoglobulin molecules in certain conditions (unpublished results of P. Guglielmi and J. L. Preud'homme).

Another point of interest is that human pre-B leukaemic cells in a fair percentage of patients¹³, as well as several Abelson virus-induced tumours¹⁵ and normal pre-B cells (G. Janossy, personal communications), were shown to contain low but detectable levels of terminal deoxynucleotidyl transferase. This enzyme was initially considered as characteristic for thymic and pre-thymic cells but it is probably present in progenitors of all lymphoid cells, a finding which might be related to the possible role of this enzyme in the generation of somatic mutations¹⁶.

The availability of tumour cells and continuous lines corresponding to early precursors of the B lineage lymphoid cells should be useful for the study of immunoglobulin gene diversification and expression. □

M. Seligmann is Professor of Immunology at the University of Paris and Head of the INSERM Unit of Immunochimistry and Immunopathology at the Hospital Saint Louis.

1. Raff, M. C., Megson, M., Owen J. J. T. & Cooper, M. D. *Nature* **259**, 224-226 (1976).
2. Melchers, F., Andersson, J. & Phillips, R. A. *Cold Spring Harbor Symp. quant. Biol.* **41**, 147-150 (1977).
3. Gathings, W. E., Lawton, A. R. & Cooper, M. D. *Europ. J. Immun.* **7**, 804-810 (1977).
4. Pearl, E. R. *et al. J. Immun.* **120**, 1169-1175 (1978).
5. Owen, J. J. T., Wright, D. E., Habu, S., Raff, M. C. & Cooper, M. D. *J. Immun.* **118**, 2067-2071 (1977).
6. Hayward, A. R., Simons, M. A., Lawton, A. R., Mage, R. G. & Cooper, M. D. *J. exp. Med.* **148**, 1367-1377 (1978).
7. Kubagawa, H., Vogler, L., Conrad, M., Lawton, A. & Cooper, M. *Fed. Proc.* **37**, 1765 (1978).
8. Abelson, H. T. & Rabstein, L. S. *Cancer Res.* **30**, 2213-2222 (1970).
9. Boss, M., Greaves, M. & Teich, N. *Nature* **278**, 551-553 (1979).
10. Siden, E. J., Baltimore, D., Clark, D. & Rosenberg, N. E. *Cell* **16**, 389-396 (1979).
11. Preud'homme, J. L., Gourdin, M. F., Reyes, F. & Fellous, M. in *Human Lymphocyte Differentiation: its Application to Cancer* (ed. Serrou, B. & Rosenfeld, C.) 345-352 (North Holland, Amsterdam, 1978).
12. Vogler, L. B. *et al. New Engl. J. Med.* **298**, 872-878 (1978).
13. Brouet, J. C. *et al. Blood* (in the press).
14. Cooper, M. D., Burrows, P. D. & Lawton, A. R. in *The Immune System: Functions and Therapy of Dysfunction* (Academic, New York, in the press).
15. Silverstone, A. E., Rosenberg, N. & Baltimore, D. in *Cold Spring Harbor Conf. Cell Proliferation*, **5** (eds Clarkson, B., Till, J. E. & Marks, P. A.) 432-452 (Cold Spring Harbor Press, New York, 1978).
16. Baltimore, D. *Nature* **248**, 409-411 (1974).

Hepatitis A: a giant leap forward

from Arie J. Zuckerman

THE past decade has seen rapid progress in the identification of viruses causing hepatitis in man, starting in 1968 with the discovery of the association between Australia antigen and hepatitis type B (serum hepatitis). Five years later, hepatitis A virus, which is the cause of infectious or epidemic hepatitis, was identified by electron microscopy in faecal extracts, and more recently evidence has been obtained of a third type of hepatitis (so called non-A, non-B hepatitis) which is probably caused by more than one virus. The worldwide research effort which has been devoted to the problem of viral hepatitis by epidemiologists, virologists, molecular biologists, immunologists, pathologists, clinicians, blood transfusion services, public health authorities and others reflects the impact which these viruses have on human health and welfare. Although remarkable advances have been made in the characterisation of the biophysical and biochemical properties of hepatitis A and B viruses, reproducible propagation of these viruses in tissue culture has defied every effort and continued to elude even the most stubborn research worker. Provost and Hilleman (*Proc. Soc. exp. Biol. Med.* **160**, 213; 1979) now report the first reliable serial cultivation of human hepatitis A virus in primary explant liver cell cultures of marmosets and in a fetal rhesus kidney cell line. This work represents a spectacular advance in a hitherto frustrating and seemingly insoluble problem.

The events which led to this success started in 1967 when Deinhardt and his associates (*J. exp. Med.* **125**, 673; 1967) transmitted human hepatitis A virus to two species of marmosets, which are small South American monkeys, and the infection was transferred serially from animal to animal. Differences in susceptibility to hepatitis A exist between the marmoset species, with *Saguinus mystax* being the most susceptible. In 1973, the CR 326 strain of hepatitis A virus was isolated in *S. mystax* by Mascoli *et al.* (*Proc. Soc. exp. Biol. Med.* **142**, 276; 1973) from naturally occurring outbreaks in Costa Rica, and the virus was further characterised in the sera and livers of

the marmosets (Provost *et al. Proc. Soc. exp. Biol. Med.* **148**, 532; 1975). By then sensitive serological tests for hepatitis A virus and its antibody became available and the susceptibility of chimpanzees as well as the susceptibility of a more readily available rufiventer-like marmoset designated as *S. labiatus* (also known as *Marikina labiata*, *Jacchus rufiventer*) was described (Provost *et al. Proc. Soc. exp. Biol. Med.* **155**, 283; 1977).

Using the CR 326 strain of human hepatitis A virus passaged five times in *S. mystax* and 26 times in *S. labiatus*, Provost and Hilleman inoculated normal liver explant cell cultures obtained from *S. labiatus* with the virus, and observed by direct immunofluorescence minute fluorescing cytoplasmic granules in some of the hepatocyte-like epithelial cell outgrowths as early as 8 days after inoculation. The number of the affected hepatocyte-like cells and the number of fluorescent granules increased with time so that when cultures were collected on day 23, 75–100 % of these cells, but not other cell types, contained large numbers of granules. There were no cytopathic changes. The virus was identified as hepatitis A by various techniques including specific immunofluorescence blocking, serum neutralisation, immune adherence haemagglutination and radioimmunoassay; and the virus was passaged serially five times. Because the availability of marmosets is very limited, an alternative cell culture system was sought. The virus propagated in primary kidney cell cultures of *Cercopithecus aethiops* (grivet or green monkey), but many of the cultures contained cytopathic adventitious agents of endogenous origin. Consequently, Provost and Hilleman used cultures of fetal rhesus kidney cells (designated FRhK6 cell line), which are epithelial-like but have only a maximum population doubling of 12. Specific immunofluorescence was detected in these cells as early as 3 days after inoculation and occurred throughout the cell sheets during the following weeks. Cytopathic changes were not observed. The virus was passaged serially, the incubation period for development of fluorescence was shortened, and there was a larger yield in these cells than in the liver cell explants. Titration of infectivity of the virus collected in the fifth and sixth passage yielded titres of 10^6 – 10^7 infectious units per ml respectively, and virus in the fifth passage collection retained the ability to infect *S. labiatus* marmosets. The virus was successfully transferred through eight serial passages in which the calculated dilution of the original inoculum was 2.9×10^{-26} .

The importance and practical significance of this work can be summarised as follows: the means may now be available for detecting human hepatitis A virus *in vitro*, for preparing large quantities of viral antigen for serological tests and for providing a source of viral antigen for the preparation of vaccines against an important infection which is endemic in all parts of the world.

However, answers to several questions are still required. Can the virus be cultivated from naturally occurring strains without adaptation by serial passage through susceptible marmosets and what is the minimum period of adaptation required; can hepatitis A virus be propagated in more readily available cell lines with longer population doubling times? Provost and Hilleman provide a hint that this may be so because the virus replicated to a limited extent in human diploid lung fibroblasts, which are commonly used for the preparation of vaccines. The way may thus be open in the near future to the effective control of epidemic hepatitis. □

Vacancies in nickel-aluminium and other alloys

from Robert W. Cahn

IN 1937 the metallurgist A. Taylor made a celebrated study of the structure of nickel-aluminium alloys; this called for great experimental precision, and in 1972 he returned to the problem and examined it with even greater scruple (Taylor and Doyle, *J. appl. Cryst.* **5**, 201; 1972). This second study may be said to have launched a new wave of research on the Ni–Al system and its analogues, even though few of the most recent protagonists seem to be aware of Taylor's second study.

Taylor's experiments centred on the alloy β -NiAl, which adopts the very simple B2 (CsCl-type) structure: nickel atoms at cube corners, aluminium at cube centres. Taylor measured macroscopic densities and lattice parameters as a function of composition, either side of the 50/50 composition: by putting together the two sets of measurements, he was able to deduce the destinations of excess nickel or aluminium atoms. On the nickel-rich side, the extra nickel atoms substitute in the usual way for aluminium atoms on the aluminium sublattice. The aluminium-rich side, however, is quite different: no aluminium substitutes on the nickel sublattice, instead nickel atoms disappear from the nickel sublattice, leaving nickel vacancies. For instance, at 45

Arie J. Zuckerman is Professor and Director of the Department of Medical Microbiology and the WHO Collaborating Centre for Reference and Research on Viral Hepatitis at the London School of Hygiene and Tropical Medicine, University of London.

at. % Ni, 18 % of the nickel sites are vacant, all of the aluminium sites are filled. Such vacancy populations, determined by composition and not temperature, are distinguished as constitutional vacancies.

Several other studies showed that stoichiometric β -NiAl quenched from a high temperature (as opposed to that slowly cooled) contained a high concentration of thermal vacancies; the most recently cited figure is 1.08 % of vacancies at 1,600 °C. This is a very much larger thermal vacancy concentration than is found in other metals or alloys, even just below the melting temperature; so large that on cooling the vacancies will separate out into a population of voids visible in the electron microscope (Epperson *et al. Phil. Mag. A38*, 529; 1978). 50/50 NiAl containing such vacancies, all on the nickel sublattice, must also contain substitutional defects—that is some nickel atoms in the aluminium sublattice, also called nickel antistructure atoms—to preserve the overall chemical composition: specifically, two vacancies must be accompanied by one substitutional defect. Such a trio of linked defects is now termed a triple defect and much recent research has centred on this entity.

In recent years, investigators in the US, England and Germany began to look for parallels to this curious behaviour in alloy systems isomorphous with β -NiAl, and attention soon centred on NiGa and CoGa. It became clear that these two phases always contain triple defects, even when they have been cooled slowly to eliminate thermal vacancies as far as possible. Though there has been some confusion in the literature on this point, it is now adequately clear that the triple defects in these alloys at 50/50 composition are indeed wholly of thermal, not constitutional origin; it is impossible to eliminate all thermal defects, even by very slow cooling.

A study of NiGa by Donaldson and Rawlings (*Acta Met.* **24**, 811; 1976) again made use of measured densities and lattice parameters to assess just how the concentration of nickel vacancies and gallium antistructure atoms varies with composition. To obtain these values, they developed an iterative method of calculation, starting out with an arbitrary assumption about the number of triple defects at stoichiometry, and refining this value in cycles until experimental densities and lattice parameters at a range of compositions could be matched. (To do this, it is necessary to assume a value for the relaxed volume of a nickel vacancy, which has to be based on theoretical

estimates, and comparison of some recent papers shows that this is the weak link in the chain of theoretical deduction.) Donaldson and Rawlings found just over 3 % of vacancies and half that number of antistructure atoms in 50/50 NiGa. At 47 % Ni, there are 7 % vacancies and 0.2 % antistructure atoms, while at 54 % Ni, the defect numbers are 1 % and 4.5 % respectively. The triple defect itself (a 2/1 ratio of the two kinds of defect) only exists at stoichiometry. Ho *et al. (Scripta Met.* **11**, 1159; 1977) obtained rather similar values, though they do not match Donaldson's and Rawlings' figures in every particular.

Berner (dissertation, Stuttgart, 1976; Berner *et al. J. Phys. Chem. Solids* **36**, 221; 1975) has found closely similar values for defect densities in CoGa.

The computational problem of determining defect concentrations from experimental data when both defect types are present has now been eased by an ingenious application of statistical thermodynamics to the problem. Edelin (*Acta Met.* **27**, 455; 1979) calculated the relations between composition, defect concentrations and defect formation energies, using merely the hypothesis that the formation energies for Ga vacancies and antistructure atoms on the Ni (or Co) sublattice are so much higher than the energies for Ni (or Co) vacancies and antistructure atoms on the Ga sublattice that the former two defects never appear in detectable numbers. Edelin showed that the variation of concentration of the two defect types, both as a function of temperature and as a function of composition, was determined by a single composite activation energy, Q , and Q can be determined directly by fitting, for a range of compositions, experimental values of density and lattice parameter to his theoretical equations. Further, it appears that when fitting is done in this way, the relaxed volumes of vacancies and of antistructure atoms can be deduced from the measurements and do not need to be used as assumed input values. (This claim seems so remarkable that it deserves critical appraisal by a suitably qualified theoretician: it would seem to be the first claimed method for deductively obtaining relaxed defect volumes from experimental data.)

The claim by Edelin to be able to measure the relaxed volumes of defect sites is of particular interest because it has been established, from diffuse X-ray scattering measurements made at Stuttgart, that the lattice near a vacancy in NiAl or CoGa is severely distorted in a [111] direction. This distortion is of the same type as is found in a $\beta \rightarrow \omega$ phase transformation. Such distortions must surely affect the relaxed vacancy volume (Ortiz and Ep-

person, *Scripta Met.* **13**, 237; 1979, for NiAl; Kirchgartner and Gerold, *J. appl. Cryst.* **11**, 153; 1978, for CoGa).

The composite activation energy, Q , found by Edelin for NiGa is 0.29–0.32 eV per atom and for CoGa, 0.265 eV per atom. These very low values account for the high thermal defect concentrations found even at stoichiometry in these two alloys. Edelin's method has not been used for NiAl (no accurate measurements of thermal defect concentrations at stoichiometry being available), but Q can be estimated from measurements of nickel diffusion in NiAl at a range of compositions (Hancock and McDonnell *Phys. Stat. Sol. A4*, 143; 1971). On the low-nickel side, it is assumed that the measured activation energy for diffusion is purely that for vacancy migration (because there is an abundance of constitutional vacancies available); then it is possible to compute the formation energy for thermal vacancies by subtracting the migration energy from the measured activation energy for diffusion at stoichiometry (where no constitutional vacancies exist). This approach gives 0.70 eV per atom, corresponding to the much smaller thermal vacancy concentration in this alloy. This much higher Q value also accounts, by way of Edelin's theory, for the virtual absence of aluminium antistructure atoms on the nickel-poor side in this alloy series, as contrasted with NiGa.

There have been several attempts to bypass the density/lattice parameter approach, which makes severe demands on experimental accuracy. (One difficult problem is to get precise densities in view of the presence of casting porosity in ingots; Taylor and Doyle overcame this by high-pressure annealing of ingots, and other investigators have sought to bypass it by using dilatometry instead of Archimedean weighing.) An alternative approach is to measure thermodynamic activities at high temperature and to relate such activities to defect concentrations. A theory to make this possible has been constructed by Neumann, Chang and Lee (*Acta Met.* **24**, 593; 1976) and has been applied to NiGa (Neumann *Scripta Met.* **11**, 969; 1977). The vacancy concentrations in NiGa alloys deduced from activity values are considerably lower than those obtained by the conventional method with quenched samples, but a very recent paper (Neumann and Chang *Z. Metallkde.* **70**, 118; 1979) shows that if allowance is made for loss of vacancies during quenching, then agreement is better. Nevertheless, it does seem that there must be some residual error in Neumann's statistical thermodynamics, and one possible source of such error is the tacit assumption that vacancies and antistructure atoms are randomly

disposed. Delavignette, Richel and Amelinckx (*Phys. Stat. Sol. A* **13**, 545; 1972) have adduced electron-microscopical evidence pointing to such order among the vacancies in hypostoichiometric NiAl, and Reynaud (*J. appl. Cryst.* **9**, 263; 1972) has evidence suggesting ordered antistructure atoms in hyperstoichiometric NiAl. Any short-range order among defects would have a considerable effect on thermodynamic activities. The recently observed anisotropic distortion around vacancies, already cited, with its associated stress field, could also affect activities. The activity method is therefore suspect, but it must be admitted that substantial order or lattice distortions could introduce errors into the density/lattice parameter method as well.

Surprisingly little attention has been paid to the question why β -NiAl, NiGa

and CoGa form defect structures at all. For instance, Massalski and Mizutani, in their survey of electronic structures of Hume-Rothery phases (*Prog. Mater. Sci.* **22**, parts 3/4; 1978) make no reference to these phases. Taylor and Doyle (*loc. cit.*) point out that the electron/atom ratio remains strictly constant at 3/2 from 50/50 NiAl to the phase boundary on the aluminium-rich side (where there are only 1.75 atoms per cell), and they assume this feature to govern the introduction of vacancies. The question then arises why other isomorphous alloy series such as Ni-Zn or Co-Be do not behave in this way. Machlin (*Scripta Met.* **13**, 123; 1979) has made an ingenious approach to this question by suggesting that the self-energy of the B (non-transition) component may be lowered by a reduction in the number of bonds between

B and transition metal atoms, resulting from the presence of transition metal vacancies adjacent to some B atoms. He is able to show that on this hypothesis, constitutional vacancies should be stabilised only for trivalent B metals, such as Al or Ga, and this is in accord with observation. There is scope for further theoretical treatment of the reasons for the formation of constitutional vacancies.

Clearly NiGa and CoGa in particular are ideally suited as test materials in which the predictions of precise statistical thermodynamics applied to defect populations can be compared with experimental observations. They are probably the best metallic materials for such comparisons, free of the complications arising from the need for charge neutrality which are unavoidable with ionic crystals. $\square$

Interferons and inducers *in vivo*

from Nowell Stebbing

MUCH is known about the mode of action of the interferons at the molecular level (Friedman *Bact. Rev.* **3**, 543; 1977), and intensive effort is being devoted to investigations of their structure and the cloning of the requisite genes. Although less attention has focused on the factors necessary for interferons to be effective in the treatment of virus infections, reports are now beginning to demonstrate the importance of host factors. The most recent such report (Shellekens *et al. Nature* **278**, 742; 1979) shows that interferon (from human leukocytes) is protective in monkeys against infection with a strain of vaccinia virus that does not respond to interferon in cultured cells. (The same cell lines are protected by the interferon against infection with vesicular stomatitis virus.) These results indicate that, *in vivo*, leukocyte interferon at least, has anti-viral effects by means other than the known molecular mechanisms. In view of the diversity of interferons in terms of immunological differences and cellular origin, it will be important to determine whether other interferons behave similarly, before these observations can be generalised. However, it is possible that the known molecular effects of interferons do not contribute directly to their anti-viral effects *in vivo* in any situation. Even when an interferon does inhibit virus replication in cell cultures, anti-viral effects in whole animals may be due entirely or in part to properties of particular cells of the host. Clearly, the anti-viral effects of interferons against infections of fibroblasts in culture should not be

extrapolated to treatments of whole animals, although this has been done for many years.

The issue of whether interferon inducers have anti-viral effects by means other than or additional to known molecular mechanisms has been highly contentious. Non-interferon-mediated anti-viral effects of polynucleotides have always been a possibility and claims that interferon induction is the sole mechanism of action have generally ignored the fact that such claims require quantification: it is necessary to know the extent to which observed anti-viral effects are due to interferon induction. It is now apparent that quantification by extrapolation from tissue culture studies is misleading. The most recent evidence in favour of interferon as the sole mediator of the anti-viral and anti-tumour effects of the double-stranded polynucleotide interferon inducer, poly(I):poly(C), is that of Gresser *et al. (Int. J. Cancer* **21**, 72; 1978). They demonstrated that anti-mouse interferon serum raised in sheep abolished the anti-viral and anti-tumour effects of poly(I):poly(C) administered to mice. However, mice treated with an anti-interferon serum, without simultaneous poly(I):poly(C) treatment, died earlier from virus infection than control mice and therefore the anti-serum may primarily affect virus- rather than poly(I):poly(C)-induced interferon. For such reasons it remains possible that poly(I):poly(C) has anti-viral and anti-tumour effects in addition to interferon induction.

Depletion of macrophages in mice obliterates the anti-viral effects of crude (macrophage) interferon and decreases the anti-viral effects of purified fibroblast interferon (Stebbing *et al. Inf.*

Immun. **19**, 5; 1978). Moreover, interferon and interferon inducers (Gidlund *et al. Nature* **273**, 759; 1978; Herberman *et al. Cancer Treatment Repts* **62**, 1893; 1978) are known to increase natural killer cell activity, implying that activities of these cells may be important for *in vivo* efficacy of interferons. Thus particular cell populations in whole animals seem to be decisive in determining the anti-viral and anti-tumour effects of interferons, and further work on the nature of their involvement would seem important for several reasons. It will of course be fascinating, to determine whether and in what ways the known molecular and cellular effects of interferons contribute to anti-viral and anti-tumour effects *in vivo*. In addition, clinical potential can only be determined by clinical studies whatever the theoretical arguments. However, it seems likely that the natural or induced immunosuppression associated with particular diseases could limit the efficacy of interferon therapy. The disappointment of negative results in clinical trials might be avoided by considering the state of cell populations necessary for *in vivo* efficacy of the interferon being tested. There is also the possibility of devising means of directly eliciting cell mediated anti-viral effects by short-circuiting the interferon mechanism.

Although interferon induction has generally been considered the means whereby polynucleotides have anti-viral and anti-tumour activity, other means are now apparent whereby even double-stranded polynucleotides could inhibit virus replication, for example by direct inhibition of protein synthesis (Levin & London *Proc. natn. Acad. Sci. U.S.A.* **75**, 1121; 1978; Kerr &

Nowell Stebbing is Director of Biology at the Research Division of Searle, High Wycombe, UK.

Brown *Proc. natn. Acad. Sci. U.S.A.* **75**, 256; 1978) possibly by means of a nuclease which degrades viral mRNA (Baglioni *et al. Nature* **273**, 684; 1978). The last possibility provides a mechanism for interfering with virus replication in a manner related to that of interferons and their inducers but without actual production of interferon. The anti-viral effects of single-stranded polynucleotides which do not induce interferon may occur by this means or by mimicking strategic nucleotide sequences in viral RNA (Stebbing *et al. Ann. N.Y. Acad. Sci.* **284**, 682; 1977). It is noteworthy that early studies on interferon induction by polynucleotides involved largely single-stranded RNAs (Isaacs *et al. Lancet* **ii** 113; 1963; Rotem *et al. Nature* **197**, 564; 1963) and that two separate mechanisms may have been involved in the overall anti-viral effect obtained. Induction of interferon was certainly a factor but independent more direct anti-viral effects may also have been involved. □

Flattening of Uranus and Neptune

from David W. Hughes

PLANETARY flattening or oblateness, ϵ , is defined as the difference between the equatorial and polar radii divided by the equatorial radius. It is not an easy quantity to determine especially in the case of the outer planets. The flattening can be obtained geometrically by using for example a double image micrometer in the focal plane of a large telescope or by the analysis of the isophotes of the image of the planet obtained from spacecraft photographs. It can be measured dynamically by observing the orbital perturbations of the inner satellites of the planet induced by its unsymmetrical mass distribution, calculating J_2 , the quadrupole moment of the gravitational field and using the simple relationship between this quantity and the flattening ϵ . Unfortunately if we take the fairly typical case of Uranus, the inner satellites Miranda, Ariel and Umbriel form a system with near commensurability of mean motions, a state of affairs that results in important mutual perturbations.

A. H. Cook of the Cavendish Laboratory, University of Cambridge reviews recent observations of ϵ in the light of the new determinations of the spin periods of the planets and shows that some of the values obtained for ϵ are implausible to say the least. His conclusions are presented in a recent edition of the *Monthly Notices of the Royal Astronomical Society* (**187**, 39P;

David W. Hughes is a Lecturer in the Department of Physics at the University of Sheffield.

1979).

For a planet in hydrostatic equilibrium the product C/Ma^2 , where C is the moment of inertia about the polar axis, M is the mass and a is the radius, must be between specific limits. This quantity is 0.4 if the planet is of uniform density and drops to 0 for an extreme central condensation.

Unfortunately the present day values obtained for the dynamical properties of Uranus and Neptune are confusingly diverse. For Uranus, recent measurements of the spin period give two inconsistent values, one about 24 h and the other about 14.5 h. There are two values for Neptune as well, one about 22 h and the other about 18.5 h. Inconsistent values have been found for one or both planets for ϵ and J_2 .

Cook then considers all possible combinations of the observed parameters to see which if any can be rejected due to the physical unreality of the value of C/Ma so produced. The central pressures of Uranus and Neptune are comparable with those of the Earth, however the mean molecular weights of the material in these planets are less and the material is more compressible. Cook concludes that these planets have C/Ma values which are definitely greater than 0.25 and probably in excess of 0.3. On this basis both the values obtained for the geometrical flattening of Uranus, 0.028 and 0.01, are inadmissible. It is not possible to decide between the three J_2 values of 0.00343, 0.005 and 0.012. For Neptune the ϵ and J_2 values both support the lower spin period of 18.5 h.

To resolve this sorry state of affairs the astronomical community must be stimulated into taking new observations. Unfortunately, as is so often the case, the best place to take these observations is outside the distorting atmosphere of planet Earth. □

More spots before the eyes

from E. G. Richards

TREATMENT of *Escherichia coli* DNA with *EcoR*I restriction enzyme results in an estimated 500–600 fragments of DNA differing first in their molecular size and second in their base composition and sequence. If such a digest is examined by electrophoresis in agarose, a technique which separates DNA according to molecular size, a relatively featureless pattern is obtained, for the resolving power of the method is insufficient to resolve such a large number of components. How useful it would be, however, if this were possible.

Fischer and Lerman (*Cell* **16**, 191;

1979) have recently described a two-dimensional electrophoretic method which approaches this ideal. They start with the observation that a mixture of urea and formamide—both non-ionic substances—can denature DNA. That is to say, if the concentration of these denaturants is increased, a point is reached at which the DNA wholly or completely 'melts'. This critical concentration depends on the temperature, occurring at a lower concentration the higher the temperature, and also on the base composition and sequence, albeit in a complicated way.

In their first experiment Fischer and Lerman applied a sample of an *EcoR*I digest of λ phage DNA, containing six fragments, along the top of a slab polyacrylamide gel and subjected it to electrophoresis downwards. They had, however, arranged that the gel contained a gradient of concentration of urea and formamide in a direction perpendicular to the direction of migration. After electrophoresis they observed that in the part of the gel where denaturant concentration was low the six fragments were resolved according to molecular size and moved with a velocity almost independent of denaturant concentration. On the other hand, on the high-concentration side of the gel the components moved with a velocity reduced by a factor of 20 or 30 and remained near the top of the gel. At intermediate parts of the gel where the denaturant concentration took intermediate values, the components moved with intermediate velocities. The result was that the zone of DNA corresponding to each fragment was distorted into a sigmoidal curve resembling a 'melting curve'. Corresponding to each fragment there was a critical denaturant concentration at which the mobility fell to a low value and at which partial or complete denaturation took place. This critical concentration varied from one fragment to another and no doubt depended on the composition and sequence of the fragment. If the temperature at which the gel was maintained was reduced (from 60 °C to 45 °C) this critical concentration increased as expected. It may be noted that this simple description was slightly complicated for one or two fragments by the presence of more than one step in the 'melting curve', a phenomenon probably associated with the denaturation of separate domains of the DNA at different denaturant concentrations.

These observations are certainly entertaining but not of great apparent importance. This importance, however, is provided in the next and crucial stage of Fischer and Lerman's work when they rotated the denaturant gradient through 90° so that the concentration of urea and formamide increased in the direction of migration. They then ob-

served that each fragment set out on its passage down the gel with a velocity corresponding to its mobility in the native state. Sooner or later, however, each fragment reached a point in the gel at which it denatured and its mobility was reduced to a low value. At that point it came (almost) to a halt and was, as an added bonus, sharpened. A fragment came to a halt at a point which was not correlated with its size so that zones which were initially faster might be overtaken by larger and slower fragments if these were the more stable. The result was that the fragments were sorted out into the order of their stability rather than of their size.

The final step was then to perform an initial separation of the fragments in agarose to obtain a separation on the basis of size and then to lay a strip of the resulting agarose gel on top of a slab polyacrylamide gel and subject the separated fragments to electrophoresis in a direction perpendicular to the first and into a denaturant gradient. The result was six well separated spots distributed over the area of the slab. That is to say they were distributed in the two-dimensional electropherogram according to molecular size in one direction and according to stability in the other.

They then applied the method to an *EcoRI* digest of *E. coli* DNA and resolved up to 350 discernible spots. The authors suggest that further improvements in the sensitivity of detection might result in the appearance of more spots and point out that their method is capable of considerable further refinement. Meanwhile the significance of the method is demonstrated by the detection of four spots unique to a digest of lysogenic λ lysogen of *E. coli* which were not observed in uninfected *E. coli* DNA. Two of these corresponded in position to the spots obtained with λ phage DNA alone.

Meanwhile a rather similar trick was used by Creighton (*J. molec. Biol.* **128**, 235; 1979) in a study of protein denaturation. He subjected various proteins to electrophoresis perpendicular to a urea gradient in a slab polyacrylamide gel. He observed that although the mobility of these changed slowly as the concentration of denaturant was increased, possibly due to the effect of the urea on the viscosity of the interstitial fluid in which the proteins migrated, they underwent a more pronounced decrease in mobility at a critical concentration of denaturant at which denaturation took place. Creighton gives a mathematical appendix in which he analyses the distribution of protein in these sigmoid shaped zones

in terms of the mobilities and rates of denaturation and renaturation of the protein. Using this technique, he obtains further evidence of the all or none character of the denaturation of several proteins and in others of more complicated occurrences.

These two papers demonstrate again the remarkable versatility of gel electrophoresis without which life would be much less exciting to molecular biologists. It makes me wonder what the next jelly trick will be. □

Pattern formation in chick limb bud

from Jonathan Slack

SOME biologists believe that a problem is as good as solved as soon as the right animal has been found with which to work, and in research on biological pattern formation the limb buds of the chick embryo have become a favourite system for study. In an organ such as this it is immediately clear that spatial patterning presents a problem separate from that of cell differentiation. This is because the limb consists of only a few cell types (muscle, cartilage, bone and connective tissue) but is built up of a complex quasi-periodic arrangement of these components in three dimensions, and the mechanism which brings about this arrangement demands explanation. A glimpse at the difficulties involved is provided however, by two recent papers in the *Journal of Embryology and Experimental Morphology*, which report attempts to test two of the most influential recent models (Wolpert, Tickle and Stampford *J. Embryol. exp. Morph.* **50**, 175; 1979; Summerbell *ibid.*, 217; 1979).

The chick limb develops on the basis both of cell lineage and of intercellular interactions. The cells which form the myofibrils of the muscles are derived clonally from the somites, and the cells which form all the other tissues are derived from the lateral plate mesoderm (Chevallier, Kieny and Mauger *J. Embryol. exp. Morph.* **41**, 245; 1977). But the spatial pattern of both muscle and non-muscle elements seems to be controlled during outgrowth of the bud by additional mechanisms which act independently along each of the three morphogenetic axes (proximodistal, from the shoulder to the fingertips; anteroposterior, from thumb to little finger, and dorsoventral, from the upper to the lower surface of the limb).

Recent work on the formation of pattern in the proximodistal axis has been dominated by the progress zone model of Summerbell, Lewis and Wolpert (*Nature* **244**, 492; 1973). According to this model, the cells within 200–300 μm of the tip of the limb bud constitute a labile region within which

some metabolic variable called the positional value rises continuously. Because the whole bud is growing, the rate of cell division being particularly high near the tip, cohorts of cells are being displaced continually from the progress zone, and when they leave the zone their positional value becomes frozen. The subsequent development of each cell then depends both on its ancestry (muscle or non-muscle) and on its positional value. This is the 'weak' form of the model. There is also a 'strong' form which asserts that the positional value of a cell in the progress zone is equal to the number of cell divisions through which it has passed since the commencement of outgrowth, and that the n th skeletal element in the final limb consists of a cohort of cells which all underwent n divisions in the progress zone.

The original experimental basis of this model involved grafting the tips of wing buds onto stumps of different ages. This led to a controversy between the laboratories of Wolpert in London (*Nature op. cit.*) and Kieny in Grenoble (Kieny and Patou R  ux' *Arch. Embryol. exp. Morph.* **41**, 137; 1977). The former group claimed that in these combinations each part developed according to its normal fate map (mosaic development) while the latter group claimed that deficiencies or excesses resulted in a reprogramming of the tissue and a re-establishment of the normal pattern (regulative development). When war between Britain and France seemed inevitable, the deadlock was broken by the dispatch of Summerbell to repeat the experiments in Grenoble. In the resulting concord (*J. Embryol. exp. Morph.* **41**, 137; 1977), both groups accepted that regulation occurs up to stage 22 but not later. At this stage of development the progress zone occupies a large proportion of the bud and so the existence of some interactions is not incompatible with the model.

This removed the strongest objection to the progress zone model, but all embryologists know that a demonstration of mosaic behaviour is not sufficient to prove a positive point because all theories of pattern formation involve a mosaic phase in between the time at which rudiments are first established and the time at which they become visibly differentiated. So a quite different type of experiment has been carried out in Wolpert's laboratory and the results are now reported (*J. Embryol. exp. Morph.* **50**, 175; 1979). Chick wing buds were subjected to a dose of X-irradiation sufficient to kill a certain proportion of their cells and the irradiated buds were allowed to develop after transplantation to healthy embryos. In the resulting limbs the proximal parts were reduced or absent

while the distal parts were well formed. The more intense the dose of radiation the greater the extent of the affected region. This is explained by arguing that the radiation kills a certain fraction of cells throughout the bud. Rudiments which are laid down shortly after irradiation are reduced in size for two reasons: first because a proportion of their cells has been killed, and second, because displacement from the progress zone is slowed down by the reduction in the proportion of dividing cells more distally. Because only the surviving cells proliferate, the progress zone eventually becomes repopulated by healthy cells and so the rudiments which leave the zone last, which are the most distal parts, develop to their normal size.

Does this work provide support for the progress zone model? It does indicate that the rudiments of the limb are laid down in proximodistal sequence, and that at a given stage the proximal part of the bud is a mosaic while the distal part is still labile, so to this extent it is compatible with the model. But it does not seem to me to have any bearing on the mechanism of generation of positional values in the rudiments. In particular it does not connect positional values with the cell cycle, for the sequence of divisions undergone by the surviving cells is presumed to be normal. So I think that it would be wrong to conclude that any evidence has been produced in favour of the strong form of the progress zone model.

Let us now turn our attention through 90° and consider the anteroposterior axis of the bud. The mechanism of pattern formation in this axis seems to be somewhat different from that in the long axis. There is a zone of tissue at the posterior edge of both the wing and the leg bud which is called the zone of polarising activity (ZPA). It is so called because when it is grafted to the anterior edge, it influences the surrounding tissues and causes them to form a second set of limb structures which are arranged in a relation of mirror symmetry to the original set (Saunders *Ann. N.Y. Acad. Sci.* **193**, 29; 1972). It has been suggested that the ZPA is the source of a morphogen that can diffuse into the other parts of the bud and is there destroyed. In the steady state situation the morphogen would be present as a gradient of concentration running from posterior to anterior and its function would be to evoke different pathways of cell differentiation at different levels (Tickle, Summerbell and Wolpert *Nature* **254**, 199; 1975). Some progress has been reported towards the chemical

isolation of the putative morphogen by McCabe and Parker (*Dev. Biol.* **45**, 349; 1975; *ibid* **54**, 297; 1976). Because this is probably the best analysed example of a morphogenetic gradient in the whole of embryology it was with some dismay that workers in the field listened to the talk given by Saunders at a meeting of the British Society of Developmental Biology in September 1976. For Saunders, who had discovered the ZPA in the first place, claimed at the meeting that it had no function at all in normal development. One argument in favour of this position was that buds whose ZPA had been removed at an early stage could go on to produce limb structures which, although often incomplete, were arranged in the normal manner (Fallon and Crosby *J. exp. Zool.* **193**, 449; 1975). The other was that tissue other than the ZPA, for example flank mesoderm or mesonephros, could show the same activity, indicating that the property is not very specific.

To counter these objections to the morphogen model, Summerbell has now reported an experiment designed to demonstrate the activity of the ZPA *in situ* without the need to graft it to an unnatural position (*J. Embryol. exp. Morph.* **50**, 217; 1979). This is done by inserting a small piece of tantalum foil into the bud at right angles to the anteroposterior axis so as to interrupt the transmission of any signal. In the limbs which develop after this operation structures are usually present only on one side of the barrier and when the barrier is placed near the centre of the bud the sequence of structures is truncated in the anteroposterior direction. The explanation for this is that the concentration of morphogen rises on the side of the barrier nearer the ZPA and falls on the side away from it. So a structure which is formed in response to a particular morphogen level will be deleted if its threshold lies in the concentration gap or will be shifted towards the barrier if it is still represented in the interrupted gradient. Summerbell produces computer simulations to demonstrate the reasonableness of this argument, and also describes a control experiment which shows that the gaps in the pattern are not simply the result of damage caused by insertion of the barrier. In many ways the logic and the results of this experiment are similar to those of constriction experiments on insect eggs (reviewed by Sander *Adv. Insect Physiol.* **12**, 125; 1976) which are also believed to indicate the existence of a morphogen gradient with a high point at the posterior end.

It would certainly be unfortunate if the ZPA did not have a role to play in normal development because it has now been detected in the limb buds of

mammals, reptiles and amphibia as well as birds and has found its way into a number of textbooks. But plausible though it is, it is difficult to see how the morphogen gradient model can be proved conclusively by macroscopic experiments alone, and whether the chick embryo limb bud will be capable of supporting the investigations down to the microscopic level of cellular physiology and biochemistry still remains to be seen. □



A hundred years ago

On the night of Sunday, May 25, loud bellowings were heard by the dwellers on the northern slopes of Etna. Towards the morning of the 26th these increased, and about midday a dense cloud of smoke was seen to issue from the side of the mountain below the great crater, apparently half way between Randazzo and Linguaglossa. This cloud increased, and on the 27th the mountain was rendered invisible, and an effect like that of an eclipse resulted. A rain of fine black ash, "like powdered emery," fell for miles around, and was so thick that Capo di Schiso could not be seen from Taormina, distance of two miles. This black rain continued all day accompanied by thundering noises from the mountain. No exact information could be procured concerning the position of the centre of disturbance, because no one could approach the new craters. During the night of the 27th the ashes continued to fall, and "huge fires could be seen looming through the black clouds"—no doubt the reflection of the molten lava on the smoke above it. It was reported in Piedemonte, a village on the north-east flanks of Etna, that three craters about a mile apart had opened at the points of a triangle, about six miles above Passo Pisciaro, a posting station nearly midway between Randazzo and Linguaglossa. Lava was said to be flowing in a valley to the north of the Val del Bove. On the 28th a great stream of lava was seen from Taormina to be descending the mountain in the direction of Randazzo, "while from the new craters great balls of fire were thrown high in the air, and burst into showers of fire like gigantic rockets, accompanied by thundering explosions." On May 29 the lava was still flowing, but the shower of ash was diminished. The facts, as above stated, were witnessed by an Englishman living in Taormina, 800 feet above the sea, at the north-eastern termination of the flanks of Etna, about fifteen geographical miles from the new craters.

From *Nature* **20**, 12 June, 158; 1979.

Jonathan Slack is a staff scientist at the Imperial Cancer Research Fund Mill Hill Laboratories.

review article

Weak interactions

Mary K. Gaillard

CERN, Geneva, Switzerland and LAPP, Annecy-le-Vieux, France

Recent progress in weak interaction physics is reviewed. The status of the field and its prospects for the future are summarised.

THE last decade has seen a revolution in the theory of weak interactions. Previously, a phenomenological prescription allowed the successful correlation of a large amount of data concerning nuclear β -decay and the decays of long lived elementary particles (muon, pion and the so-called strange particles) as well as neutrino-induced reactions and muon capture by nuclei. All these processes could be attributed to a simple interaction, the self-coupling of a single 'charged current'. However, the theory remained phenomenological in the sense that processes involving multiple weak transitions, or the interplay of weak and electromagnetic interactions, could not be calculated. In contrast, purely electromagnetic processes have long been understood; the theory permits the calculation of effects to any order in the electromagnetic coupling once the value of electric charge has been specified.

The new developments in weak interactions have resulted in the elaboration of a theory with a calculational power¹ equal to that of quantum electrodynamics. However, while measurements of electromagnetic processes have confirmed the theory to as many as six orders in the coupling constant, it is extremely difficult to test weak interaction theory beyond the lowest order. This is because the coupling is indeed weak, and higher order effects are generally beyond the limits of experimental precision. There are some measurable effects involving K-mesons which are forbidden at lowest order in the weak coupling and therefore measure directly the strength of higher order transitions. However, the interpretation of these phenomena is clouded by the effects of strong interactions.

Although they lack a laboratory for direct tests of the calculational power of the theory, most particle physicists are, nevertheless, convinced that they are on the right track. This is because the construction of the theory required the introduction of new interactions in the lowest order and of new elementary particles, neither of which had been observed. The spectacular confirmation of these predictions form the basis for the present confidence that there is at last a true theory of weak interactions. This article will deal with these new phenomena in so far as they are relevant to neutrino physics.

Neutral currents

The weak transitions observed before 1973 could all be attributed to a mechanism involving the transmission of one unit of electric charge. This concept is best visualised in terms of pairs of elementary fermions such as the electronic leptons (e^- , ν_e), the

muonic leptons (μ^- , ν_μ), the lightest quarks (u, d), and the corresponding pairs of antifermions. The members of each pair differ by one unit of electric charge, and 'charged current' processes are those in which charge is exchanged between pairs as, for example, in neutrino induced inverse muon decay:

$$\nu_\mu + e^- \rightarrow \mu^- + \nu_e \quad (1)$$

or in neutrino induced reactions involving hadrons, which are ascribed to a charge exchange reaction on quarks, the presumed constituents of hadrons:

$$\nu_\mu + d \rightarrow \mu^- + u \quad (2)$$

The neutral current counterparts of reactions (1) and (2) would be, for example:

$$\nu_\mu + e^- \rightarrow \nu_\mu + e^- \quad (3)$$

$$\left. \begin{aligned} \nu_\mu + d &\rightarrow \nu_\mu + d \\ \nu_\mu + u &\rightarrow \nu_\mu + u \end{aligned} \right\} \quad (4)$$

Reactions (3) and (4) were long believed to be absent at lowest order in the weak coupling. However, the occurrence of reaction (1) and its inverse implies that reaction (3) should also occur via a doubly weak transition

$$\nu_\mu + e^- \rightarrow \mu^- + \nu_e \rightarrow \nu_\mu + e^- \quad (5)$$

although at a much lower rate, since the rate for reaction (1) is already very small. The difficulty in constructing a real theory for weak interactions was that if one tried to calculate the rate for reaction (5) it turned out to be infinite.

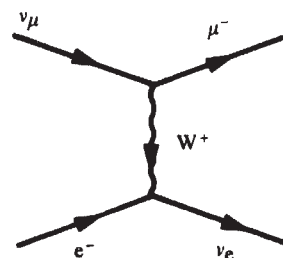


Fig. 1 Inverse muon decay as mediated by a weakly coupled massive vector boson.

The problem is that one must sum over all intermediate ($\mu^- \nu_e$) states in reaction (5), including those that violate energy conservation up to an amount allowed by the Heisenberg uncertainty relation $\Delta E < \hbar/\tau$ where τ is the interaction time. As far as we know from experimental results, the weak interaction is 'point-like'. That is, it has zero range and is instantaneous. That means that intermediate states of infinite energy contribute to reaction (5) with no damping, and the sum does not converge. This problem is avoided by introducing a finite space-time extension for the weak interaction, which is consistent with the data as long as the range introduced is shorter than distances which can be probed by present experiments. In quantum field theory the range of an interaction is governed by the mass of the mediator field: in quantum electrodynamics the mediator is the massless photon and hence the interaction range is infinite. For weak interactions the mediator must be sufficiently massive so that the effects of a non-zero range ($\tau \sim r/c \sim \hbar/mc^2$) are beyond the reach of present experiments. For the conventional charged current interactions, Fig. 1, the presumed mediator is a charged vector boson $W^\pm$, and its mass must be greater than about $20 \text{ GeV } c^{-2}$.

The introduction of the charged bosons $W^\pm$ suffices to damp high energy contributions ($\Delta E \geq m_W c^2$) to the sum over intermediate states in reaction (5). At the same time we acquire new weak processes like the production of a pair of charged bosons via the exchange of a lepton, Fig. 2:

$$\nu_\mu + \bar{\nu}_\mu \rightarrow W^+ + W^- \quad (6)$$

While the W s must be so massive that their production would be difficult at present energies, we can consider higher order processes like

$$\nu_\mu + \bar{\nu}_\mu \rightarrow W^+ + W^- \rightarrow \nu_\mu + \bar{\nu}_\mu \quad (7)$$

again involving a sum over all energies of the intermediate W - W system. Now a new difficulty arises. Experimental studies of weak processes have demonstrated that the intermediate boson must carry one unit of intrinsic angular momentum (spin), like the photon. However, unlike the photon the W is massive and therefore has three degrees of freedom instead of two. The amplitude for production of longitudinally polarised W s grows with energy like $E/m_W c^2$, so that the sum over intermediate states is not damped in spite of the space-time extension of the interaction in Fig. 2.

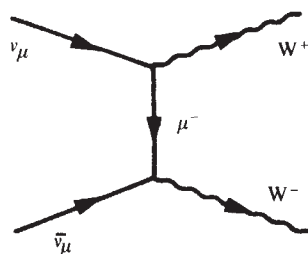


Fig. 2 Charged current muon exchange mechanism for $\nu_\mu \bar{\nu}_\mu \rightarrow W^+ W^-$.

We are left with a situation where a manageable theory of weak interactions cannot be constructed without introducing new processes which can interfere destructively with that of Fig. 2, so as to eliminate the high energy contributions to the sum in reaction (7). The simplest possibility²⁻⁴—and the one apparently chosen by nature—involves the introduction of a massive, electrically neutral spin-1 intermediate boson, the Z^0 , which couples both to the elementary fermions (quarks and leptons) and also to the charged intermediate bosons. Then the mechanism of Fig. 3 can interfere destructively with that of Fig. 2, provided the coupling constants are suitably adjusted. Similar analyses of scattering processes involving charged fermions show that they

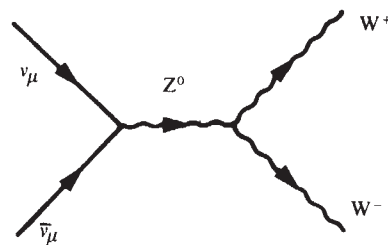


Fig. 3 Neutral current Z^0 exchange diagram for $\nu_\mu \bar{\nu}_\mu \rightarrow W^+ W^-$ which can cancel Fig. 2 to damp the reaction at high energy.

must also couple to the Z^0 , but with a coupling strength that depends on their electric charge. This is because, for example, the process

$$e^+ + e^- \rightarrow W^+ + W^-$$

can also occur via photon exchange, Fig. 4a, and it is the sum of the three mechanisms of Fig. 4 which must cancel at high energies.

The above arguments mean that the couplings of the Z^0 are tightly constrained in terms of the known charged current and electromagnetic couplings. The Z^0 couples to a fermionic current which is an admixture of a part derivable from the charged current, involving only left-spinning fermions and right-spinning antifermions (that is a $V-A$ current), and a part which is identical in structure to the (spin independent) electromagnetic current. The relative strengths of these two components depends on the relative strengths of the $W^\pm$ coupling constant g and the electromagnetic charge e . Since low energy charged current reactions measure only the effective Fermi coupling constant,

$$\frac{G_F}{\sqrt{2}} = \frac{g^2}{8m_W^2} \approx 10^{-5}/m_p^2 \quad (8)$$

(where m_p is the mass of the proton), the theory is characterised by an unknown parameter which is commonly phrased in terms of a 'weak angle' θ_w :

$$\sin \theta_w = e/g \quad (9)$$

The current coupled to the massive neutral boson Z^0 is then decomposed according to

$$J^Z = J^{(V-A)} + \sin^2 \theta_w J^{em}, \quad (10)$$

where J^{em} is the electromagnetic current and $J^{(V-A)}$ is the neutral counterpart of the charged currents. At low energies, exchange of a massive Z^0 will appear as the point like self coupling of the current (10), generating transitions like (3) and (4).

Such processes were first observed⁵ in 1973, only after their theoretical importance was realised, and provided an incentive for intensive searches. They had not been observed earlier because of the inherent experimental difficulties. Charged current transitions were extensively studied in decays long before neutrino experiments became feasible. A decay process which can occur via neutral currents, for example a nuclear transition with (e^+e^-) rather than $(e\nu)$ emission, is completely masked by the much faster electromagnetic transitions. (This is not the case for strangeness-changing decays which we will discuss later; in fact it was the failure to observe neutral current transitions in strange particle decays which led most particle physicists to assume that only charge changing weak processes were present in nature.) After about 10 years of accelerator neutrino physics, neutral current reactions remained hidden. Neutrino scattering from electrons as in reactions (1) and (3) has a tiny cross-section compared with neutrino scattering from hadrons, and without the final state muon which signs the neutrino hadron scattering process (2), the analogous neutral current process (4) remained buried under neutron-induced background reactions.

Since their discovery⁵ by the Gargamelle bubble chamber collaboration at CERN, neutrino-induced neutral current reactions have been studied in great detail⁶. The neutrino-hadron reactions which we have written in terms of the simple quark transitions (4) is translated in the laboratory to the very complex neutrino-nucleon interactions

$$\nu_\mu + N \rightarrow \nu_\mu + X \quad (11)$$

where N is a neutron or a proton and X is any allowed hadronic final state, from a single nucleon to a multi-particle system. Studies of a variety of final states have been carried out for both neutrino and anti-neutrino induced reactions. Experiments have also been performed to study the rarer $\nu_\mu e^-$ and $\bar{\nu}_\mu e^-$ elastic scattering processes, requiring the observation of an isolated electron track in a particle detector. In addition, the Z^0 exchange contribution to electron-nucleon interactions has been observed⁷ via its interference with the dominant photon exchange process. The cross-section for

$$e + N \rightarrow e + X \quad (12)$$

is found to depend on the polarisation of the incident electron, a parity violating effect which signals the contribution of a weak transition.

The simple fact that neutral current processes exist in nature does not provide evidence that cancellation mechanisms of the type described above are really at work. What is striking about the data is that they apparently confirm the precise spin- and charge-dependence of the neutral current coupling as prescribed by the requirement of such cancellations. In other words, each experiment can be analysed assuming the form of equation (10) for the neutral current, and that experiment will provide a determination of the weak angle θ_w . The different processes studied yield values for θ_w which are compatible with one another within the experimental uncertainties yielding an average value⁶ $\sin^2 \theta_w = 0.23 \pm 0.02$.

The parameter θ_w governs only the relative strengths of the components of the neutral weak current. There is a second measurable parameter which governs the overall coupling strength and is the neutral current analogue of the Fermi coupling constant, Equation (8), for charged current reactions. Constraints were obtained for the neutral current couplings by requiring cancellation of reactions at high energies where masses play no part. Those arguments only constrained the neutral current coupling relative to a particular combination of the W and photon couplings. The effective Fermi coupling for low energy neutral current reactions is therefore dependent on the Z^0 mass; explicitly one finds

$$\frac{G_F^0}{\sqrt{2}} = \frac{g^2}{8 \cos^2 \theta_w m_Z^2} \quad (13)$$

However, the neutral boson mass m_Z is not completely arbitrary. The construction of a calculable theory is more delicate than our simple discussion suggests. To ensure that the required relationships among coupling constants remain stable against all higher order corrections, scalar bosons must also be introduced, and the masses of the vector bosons depend on the scalar (Higgs particle) content of the theory. The simplest possibility^{3,4} for constructing a fully calculable theory containing three heavy vector bosons requires the introduction of a single scalar particle. In this case the vector boson masses satisfy the relation:

$$m_W^2 = m_Z^2 \cos^2 \theta_w \quad (14)$$

resulting in identical Fermi coupling strengths for charged and neutral current transitions:

$$G_F^0 = G_F \quad (15)$$

Measurements of the neutral current coupling strength satisfy this relation to within the experimental accuracy of about 5%.

To summarise: (1) neutrino experiments have revealed the existence of an interaction long thought to be absent: weak transitions with no exchange of electric charge. (2) The proper-

ties of this interaction are compatible with the requirement that higher order weak effects be calculable. (3) The experimental evidence strongly suggests that weak interactions have the simplest possible form which can satisfy the criterion of calculability.

Charm and dileptons

We have already paired the elementary fermions according to

$$(e^-, \nu_e), (\mu^-, \nu_\mu), (u, d)$$

This is not quite correct. The up quark (u) is actually paired with a superposition of the down quark (d) and strange quark (s),

$$(u, d_c); d_c \equiv \cos \theta_c d + \sin \theta_c s \quad (16)$$

where θ_c is a phenomenological parameter known as the Cabibbo angle⁸. Equation (16) means that the sum of the squared amplitudes for, say, the semi-hadronic transitions

$$d \rightarrow u + e^- \bar{\nu}_e \quad (17)$$

$$s \rightarrow u + e^- \bar{\nu}_e \quad (18)$$

equals the squared amplitude for the purely leptonic transition

$$\mu^- \rightarrow \nu_\mu + e^- \bar{\nu}_e \quad (19)$$

which is the μ -decay process. Reaction (17) is responsible for neutron and pion β -decay, while reaction (18) is responsible for the β -decay of strange particles. Comparison of the decay rates determines the Cabibbo angle: $\sin \theta_c \approx 0.2$. One of the most striking features of strange particle decays is the very strong suppression of neutral current transitions. For example, the neutral current process

$$K_L \rightarrow \mu^+ + \mu^- \quad (20)$$

is found to be suppressed relative to the charged current process

$$K^- \rightarrow \mu^- + \bar{\nu}_\mu \quad (21)$$

by a factor 10^{-8} in the decay rate.

On the other hand, if we pursue the analysis described in the previous section, we see that the higher order quark-scattering process

$$\bar{d}s \rightarrow W^+ W^- \rightarrow \bar{d}s \text{ (or } \bar{s}d) \quad (22)$$

will have an infinite amplitude unless the mechanism of Fig. 5a for $\bar{d}s \rightarrow W^+ W^-$ is cancelled by another mechanism, which one might assume to be the Z^0 exchange mechanism of Fig. 5b. However, this would imply an effective neutral current Fermi coupling strength for strangeness changing processes as strong as that for charged current ones, and in particular the decay of reaction (20) would be as rapid as the decay (21). As this is in gross disagreement with the data, the only way to render the process (22) finite is to introduce yet another process. The simplest possibility is to introduce a new quark⁹ c , with couplings to the d and s quarks such that the c -exchange diagram of Fig. 5c cancels the u -exchange diagram of Fig. 5a. This is achieved if the new quark, called a charmed quark, couples to the d, s combination which is orthogonal to the one coupled to u . Thus we introduce a new fermion doublet

$$(c, s_c); s_c \equiv \cos \theta_c s - \sin \theta_c d \quad (23)$$

To ensure that the delicate cancellations arranged for weak interactions are not destroyed by strong interaction effects, the charmed quarks must be subject to the same strong forces as the ordinary u, d and s quarks. This means that they should also form bound states. The fact that no charmed hadrons had been observed could be attributed to a higher mass for the charmed quarks and hence for their bound states. However, for this picture to be consistent with kaon data, the charmed quark mass could not be made arbitrarily high. The cancellation of the

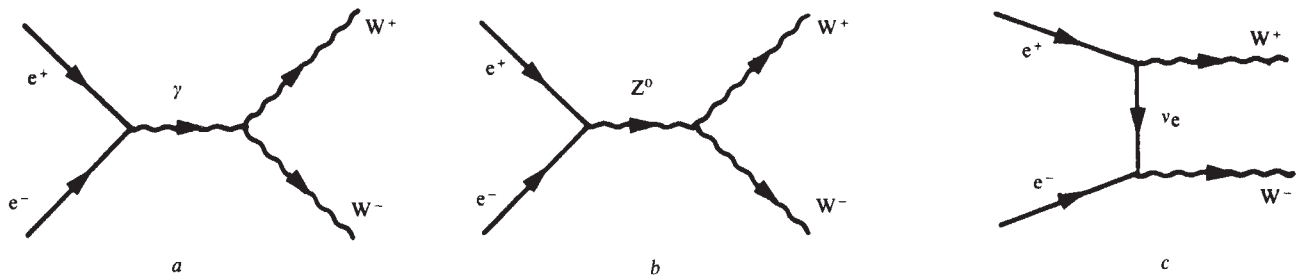


Fig. 4 Mechanisms for $e^+e^- \rightarrow W^+W^-$ which must interfere destructively at high energies in a calculable theory.

processes in Fig. 5a and c is only exact if the up and charmed quark masses are equal. Therefore, if $m_c \gg m_u$, the sum over intermediate W^+W^- states in reaction (22) will be damped only for energies $E > m_c c^2$. This is directly relevant to K decay; in the absence of a direct $Z^0 \bar{s}d$ coupling, the decay (20) can still occur via higher order mechanisms such as

$$\bar{s}d \rightarrow W^+W^- \rightarrow \mu^+\mu^-$$

which will have an effective Fermi coupling determined by the charmed quark mass:

$$G_F^{\text{eff}} \sim \cos^2 \theta_c \sin^2 \theta_c G_F^2 m_c^2 \quad (24)$$

The tiny rate observed for the decay (20) is consistent^{10,11} only with a charmed quark mass no greater than one or two $\text{GeV } c^{-2}$.

The startling discovery of the J/ψ particle with a mass of $3.1 \text{ GeV } c^{-2}$ at Brookhaven¹² and at SLAC¹³ in 1974 was immediately interpreted by several particle physicists as a charm-anticharm quark bound state. However, it took about two years before this identification was firmly established, and the study of the J/ψ family of $\bar{c}c$ states led to considerable progress in the understanding of strong interaction dynamics. We describe here, however, how the weak couplings of charmed particles were studied, particularly in neutrino experiments. Once again these couplings are completely determined^{9,14} by the requirement that the theory be calculable (cancellation between Fig. 5a and c), and they are characterised by two features: (1) the (c, s) coupling dominates the (c, d) coupling by a factor $\cot \theta_c \approx 5$ in amplitude; (2) the charged current coupling is of the usual $V-A$ type.

How are charmed particles produced in neutrino experiments? The nucleon is visualised as a system of bound quarks: three valence quarks which determine its quantum numbers (uud for the proton, ddu for the neutron) and quark-antiquark pairs which are referred to as the quark 'sea'. A neutrino incident on a nucleon can create a charmed quark via reactions like

$$\nu_\mu + d \rightarrow \mu^- + c \quad (25)$$

$$\nu_\mu + s \rightarrow \mu^- + c \quad (26)$$

where in reaction (25) the d can be a valence or a sea quark, but the strange quark in reaction (26) must come from the sea. The

weak interaction breaks up the nucleon, and the charmed quark picks up, say, a $\bar{u}$ or $\bar{d}$ antiquark from the hadronic debris to form a charmed D meson. The D meson will then decay weakly, mainly via the strangeness changing processes

$$c \rightarrow s + \{\mu^+ \nu_\mu \text{ or } e^+ \nu_e \text{ or } u\bar{d}\} \quad (27)$$

One therefore expects to find K mesons in the final state when a charmed meson has been produced. This prediction has been verified both in e^+e^- pair creation of D mesons^{15,16} (via the electromagnetic process $e^+e^- \rightarrow \gamma \rightarrow \bar{c}c$) and in neutrino production^{17,18} of a single charmed particle. When the D decays into a system of only charged hadrons, it can be identified directly by the reconstruction of its mass ($1.86 \text{ GeV } c^{-2}$) from the momenta of its decay products. If the D decays semi-leptonically some of its energy is taken by the unobserved neutrino. In this case the event is signed by an extra lepton in the final state. Such events were found⁶ as expected to be characterised by an excess of kaons.

Even without identification of the hadrons in the final state, it is possible to verify the expectations (1) and (2) simply by studying dimuon events:

$$(\nu_\mu \text{ or } \bar{\nu}_\mu) + N \rightarrow \mu^- + \mu^+ + X \quad (28)$$

where X represents the final state hadronic system for which only the total energy is measured. The important elementary transitions contributing to dimuon events are reactions (25) and (26), followed by the decay

$$c \rightarrow s + \mu^+ + \nu_\mu \quad (29)$$

and antineutrino production of anticharm from the $\bar{s}$ sea:

$$\bar{\nu}_\mu + \bar{s} \rightarrow \mu^+ + \bar{c} \quad \downarrow \quad \bar{s} + \mu^- + \bar{\nu}_\mu$$

Antineutrinos can also produce $\bar{c}$ from $\bar{d}$ antiquarks in the nucleon's sea. However, the sea contribution is small, and scattering from s and $\bar{s}$ is important only because it is favoured by the factor $\cot \theta_c$ in amplitude. There can also be a charm-anticharm component of the quark sea, but $(c\bar{c})$ pairs can be created only for a time $\tau \sim \hbar/2m_c c^2$ which is small compared

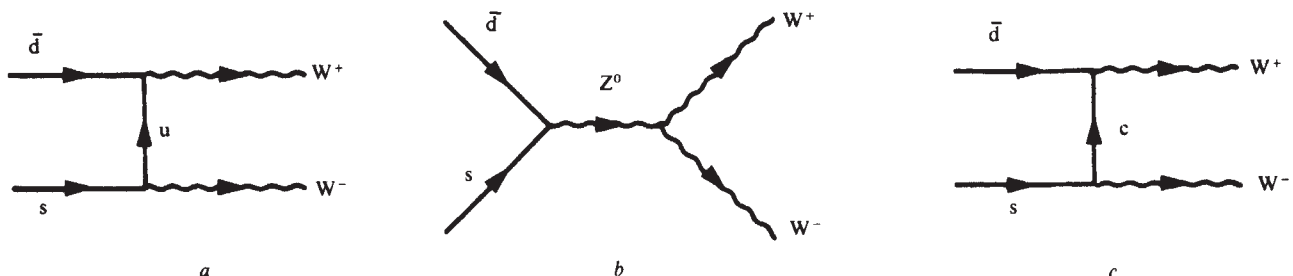


Fig. 5 Mechanisms which might interfere so as to damp the strangeness changing reaction $d+s \rightarrow W^+ + W^-$; a, Conventional charged current mechanism; b, neutral current coupling which is disallowed by K-decay data; c, charmed quark exchange.

with the interaction time unless the neutrino energy is well above the threshold for charm production.

Assuming that scattering from the charmed quark sea is unimportant, we expect the following characteristics for dimuon events.

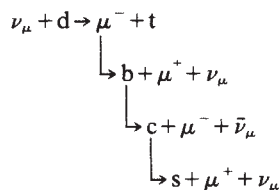
(1) Because of the larger amplitude for scattering from an s or $\bar{s}$ quark, the sea contribution is enhanced relative to that for ordinary single muon events. On average the sea quarks carry a smaller fraction of the nucleon momentum than ordinary quarks, so dimuon events should be characterised by a lower effective centre of mass energy for the scattered μ^-c or $\mu^+\bar{c}$ system. This effect should be most pronounced in $\bar{\nu}_\mu$ induced events where sea quarks make the only contribution.

(2) The final state muon angular distribution depends on the relative spin alignments of the initial interacting fermions. The distribution is isotropic in the neutrino-quark centre of mass system if their spins are both aligned (right spinning) or both anti-aligned (left spinning) with their direction of motion. This is the case for the three processes (25), (26) and (30) if the coupling to charmed quarks are of the usual $V-A$ type as predicted.

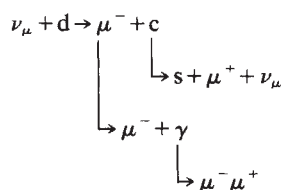
A few dimuon events, suggestive of charm production, were observed¹⁹ even before the discovery of the J/ψ . By now several high statistics studies have been performed and they support⁶ both the predictions (1) and (2) which verify, respectively, conditions (1) and (2) given previously.

Beyond charm: multileptons

The addition of a charmed quark along with u, d, and s established a lepton-quark symmetry for weak interactions: for each lepton pair there is a corresponding quark pair. This symmetry is actually necessary²⁰ if the theory is to be fully calculable; otherwise infinite amplitudes are found to occur for processes like $e^+e^- \rightarrow W^+W^-$ when higher order contributions are included. Therefore, the discovery^{21,22} of the τ -lepton and its associated neutrino led to the anticipation that still another quark pair (t, b) should exist. This anticipation received partial confirmation in the discovery of the Y with a mass of $9.4 \text{ GeV } c^{-2}$, generally interpreted as a bb bound state—leaving the t still to be discovered, presumably with a higher mass. Just as c couples preferentially to s in charged current transitions, t will couple preferentially to b, and it is not unlikely that b will decay more often to c than to u. Then cascade decays can generate spectacular multi-lepton events like



Unfortunately, the charged couplings of b and t quarks to the valence u and d quarks are constrained by other data to be very small, so that production cross-sections will not be large. Quadrimuon events have already been observed, but these data can be understood in terms of charm production accompanied by the Bremsstrahlung of a virtual photon, for example



It is hoped that beneath the background from this now 'conventional' source of multimuoons, a few events signalling still newer physics may be dug out.

Prospects for weak interactions

Assuming that no experimental surprises will upset what now seems to be the steady confirmation of a true theory of the weak interactions, what is the future of the field?

(1) A crucial test of the theory which is still out of reach of present experimentation is the establishment of the actual existence of the postulated vector bosons, $W^\pm$ and Z^0 . If the theory is correct their masses are already known from the measured values of the weak angles θ_w , the Fermi constant G_F , and the electric charge e through equations (8), (9) and (14). This gives

$$m_W = (77.8 \pm 3.4) \text{ GeV } c^{-2}, \quad m_Z = (88.7 \pm 2.7) \text{ GeV } c^{-2}$$

New facilities are being constructed at CERN and at Fermi Lab with the primary goal of producing and detecting these particles. A further important step would be their pair production by very high energy electron-positron colliding beams which would allow a probe of the WWZ coupling.

(2) The theory also requires the existence of one or more scalar particles, but little can be predicted about their masses or even their number, although recent speculations concerning the unification of strong, weak and electromagnetic interactions in terms of a single force (see below) place the mass of a scalar particle at about $10 \text{ GeV } c^{-2}$. Finding any scalar particle with the appropriate properties²³ would be an important confirmation of present theoretical ideas.

(3) Some theorists believe that detailed studies of hadrons containing the heavy b and t quarks may shed new light on the origin of CP violation, which so far has manifested itself only in the isolated neutral kaon system.

(4) The present theory contains an arbitrary parameter, the weak angle θ_w . This parameter is indeterminate except within the context of an enlarged theory containing new couplings and new vector bosons. Much theoretical effort is being devoted to the construction of a theory which not only determines θ_w , but describes the strong, weak and electromagnetic forces in terms of a unique coupling constant. There has been considerable progress in this direction, but the experimental predictions which could test such a theory are up to now extremely limited, and even higher energies do not promise much more. The most exciting possibility at present is that the proton may be unstable and have a lifetime which is accessible to experimental measurement.

(5) The outstanding problem is an almost total lack of understanding of the fermion spectrum. How many fermion pairs are there? Why are their masses so disparate? Are neutrino masses absolutely zero? Why do quarks form weakly interacting pairs according to particular superpositions of quark states? Some of these questions have been partially considered by the attempts to unify forces discussed above. Some of their answers may also be related to the scalar particle content of the theory. But the general theoretical stalemate with respect to the elementary fermions suggests most persistently that further surprises may be in store.

1. 't Hooft G. *Nucl. Phys.* **B33**, 173; **B35**, 167 (1971).
2. Glashow, S. L. *Nucl. Phys.* **22**, 579 (1961).
3. Weinberg, S. *Phys. Rev. Lett.* **19**, 1264, 1967.
4. Salam, A. *Proc. 8th Nobel Symp.* (Almqvist and Wiksells, Stockholm).
5. Haset, F. *et al. Phys. Lett.* **46B**, 121 and 138 (1973).
6. Baltay, C. *Proc. 19th Int. Conf. on High Energy Physics*, Tokyo (in the press).
7. Prescott, C. Y. *et al. Phys. Lett.* **77B**, 347 (1978).
8. Cabibbo, N. *Phys. Rev. Lett.* **10**, 531 (1963).
9. Glashow, S. L., Iliopoulos, J. & Maiani, L. *Phys. Rev.* **D2**, 1285 (1970).
10. Vainshtein, H. I. & Khrilovich, I. B. *Zh. Eksp. teor. Fiz. Pisma Red.* **18**, 141 (1973).
11. Gaillard, M. K. & Lee, B. W. *Phys. Lett.* **33**, 108 (1974).
12. Aubert, J. J. *et al. Phys. Rev. Lett.* **33**, 1404 (1974).
13. Augustin, J. E. *et al. Phys. Rev. Lett.* **33**, 1406 (1974).
14. Gaillard, M. K., Lee, B. W. & Rosner, J. R. *Revs. Mod. Phys.* **47**, 277 (1975).
15. Goldhaber, G. *et al. Phys. Rev. Lett.* **37**, 255 (1976).
16. Peruzzi, I. *et al. Phys. Rev. Lett.* **37**, 569 (1976).
17. Cazolli, R. *et al. Phys. Rev. Lett.* **34**, 1125 (1975).
18. Bleitzhau, G. *et al. Phys. Rev. Lett.* **36**, 710 (1976).
19. Rubbia, C. *Proc. 18th Int. Conf. on High Energy Physics*, 4-117 London (1974).
20. Bouchiat, C., Iliopoulos, J. & Meyer, Ph. *Phys. Lett.* **38B**, 519 (1972).
21. Perl, M. *et al. Phys. Rev. Lett.* **35**, 1489 (1975).
22. Perl, M. *Nature* **275**, 273, (1978).
23. Ellis, J., Gaillard, M. K. & Nanopoulos, D. V. *Nucl. Phys. B* **106**, 292 (1976).

articles

Mid-Mesozoic closure of Permo–Triassic Tethys and its implications

A. M. Celâl Şengör

Department of Geological Sciences, State University of New York at Albany, Albany, New York 12222

Data from the northern Tethyan domain indicate that Permo–Triassic Palaeo-Tethys closed between the late Triassic and the mid-Jurassic. This closure was caused by the collision with Laurasia of a Cimmerian continent rifted away from northern Gondwanaland during the Triassic. The Neo-Tethys may have opened partly as a back-arc basin over the Palaeo-Tethyan subduction zone and rotation of the Cimmerian continent may have been partially responsible for mid-Mesozoic block-faulting in extra-Alpine central Europe.

HSÜ^{1,2} suggested that the Tethyan Ocean, contemporaneous with the latest Palaeozoic–early Triassic Pangaea^{3–5} was obliterated mainly during the Jurassic along a suture zone that must lie north of the late Mesozoic–Cenozoic sutures of the Tethyan orogenic system. This is an attractive hypothesis because all the late Mesozoic–Cenozoic sutures within the Tethyan Orogen represent oceans (hereafter Neo-Tethys) that opened during the Triassic and later⁶ (Fig. 1). They cannot therefore mark the site of the latest Palaeozoic–early Triassic Tethys (hereafter Palaeo-Tethys; Tethys 1 of Dewey *et al.*⁵). I review evidence here from the Alpine/Himalayan orogenic system for the existence and mid-Mesozoic closure of Palaeo-Tethys and discuss some tectonic implications of this event.

Regional review

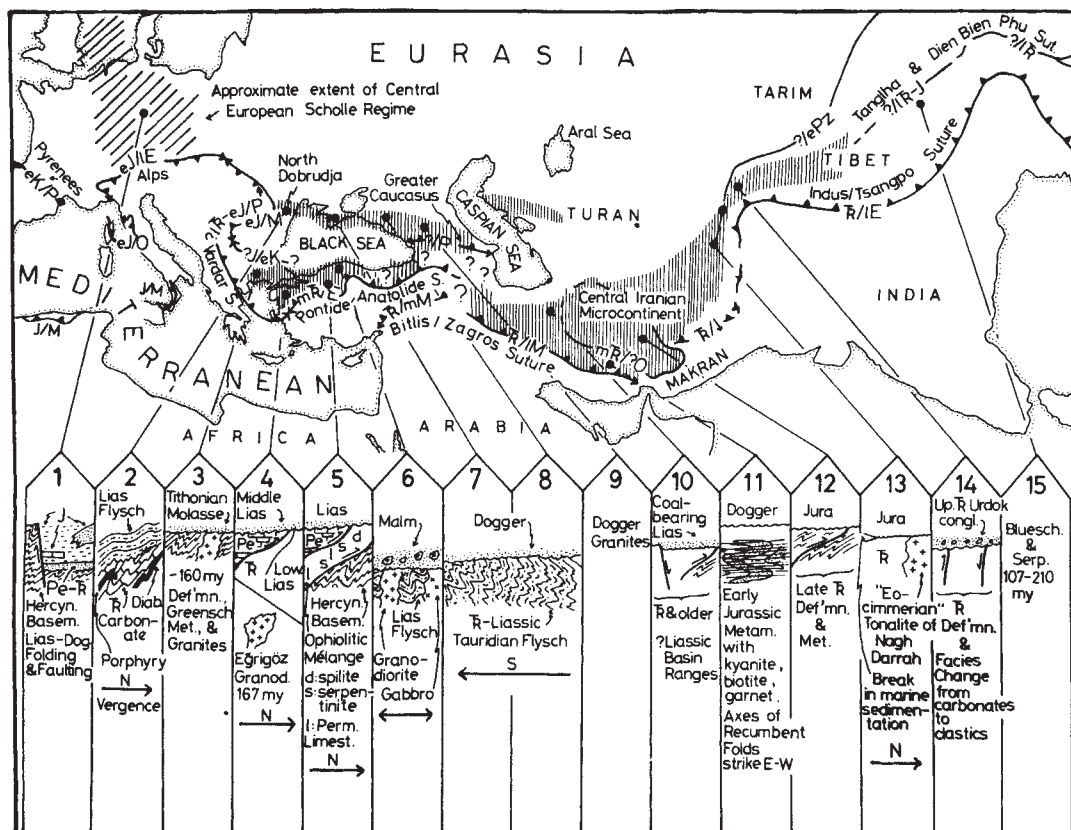
The existence of Palaeo-Tethys is implied by the observation that when the continents around the Atlantic and Indian Oceans are restored to their positions before the break-up of Permo–Triassic Pangaea^{3–8}, a large, westward narrowing triangular area appears between Laurasia and Gondwanaland. Figure 1 shows that all of the Alpine-age sutures south and west of Dobrudja–Crimea–Greater Caucasus represent oceans that opened during the Triassic or later (Pyrenees: early Cretaceous⁹; Alps/Apeninines: early Jurassic¹⁰; Carpathians/Vardar Ocean: late Triassic and/or early Jurassic⁵; Pontide/Anatolide Ocean: mid-Triassic^{11,12}; Bitlis/Zagros Ocean: Triassic¹³; North Makran/Afghanistan and Indus/Tsangpo Oceans: Triassic¹⁴). These Neo-Tethyan oceans mostly ruptured northern Gondwanaland. As far west as the Balkans they cut terrains characterised by Pan African/Baikalian or older basement, and from the Balkans westwards they cut Hercynian basement (ref. 15, Fig. 6). Evidence, both from Atlantic and Indian Ocean magnetic anomaly

lineations^{5,8} and from palaeogeographic analyses of the peri-Tethyan domain^{6,15} shows that the Permo–Triassic Palaeo-Tethys must have lain north of the Alpine sutures, separating late Palaeozoic Laurasia from late Palaeozoic Gondwanaland.

No pre-Triassic ophiolite complexes have been confirmed within the Alpine System^{1–5} and therefore it has been difficult to locate the site of the closure of Palaeo-Tethys. Because of the sparsity of ophiolites I use here a variety of criteria such as compressional deformation, Barrovian metamorphism, calc-alkaline magmatism, rapid emergence and flysch/molasse facies sedimentation collectively to identify a zone within which I suggest the Palaeo-Tethyan suture is located. One of the chief reasons why the suture zone has so far gone undetected is that it lay very close to the northern margin of Neo-Tethys and the very intense deformation associated with the closure of that complex ocean has largely overprinted the signature of the earlier collision. Figure 1 shows areas thought to have been affected by the deformation associated with the closure of Palaeo-Tethys. This area is widest where terminal Alpine collision has not yet occurred (north of Makran), an observation that supports the view that the Alpine collisions have very strongly overprinted the earlier structures.

The westernmost places where the effects of early to mid-Mesozoic convergent plate margin activity can be seen are the peri-Rhodopean Belt in the northern Aegean (Fig. 1, loc. 3; refs 16, 17) and in North Dobrudja (Fig. 1, loc. 2; ref. 18 and the bibliography therein). The precise geometry of the mid-Mesozoic deformation in the peri-Rhodopean Belt is not well-understood, but Kockel *et al.*¹⁵ have reported intense folding, greenschist metamorphism and granite intrusion sometime during the Dogger–Malm transition. The deformed assemblage is overlain by Tithonian molasse¹⁷ indicating that here deformation had long ceased when the Eo-Hellenic tectonism began in the sub-Pelagonian area in post-Tithonian times¹⁷. Whereas the peri-Rhodopean deformation may be related to the closure of Palaeo-Tethys, the Eo-Hellenic ophiolite obduction is clearly pre-collision Neo-Tethyan. In North Dobrudja the geometry is much clearer and involves a north-vergent, imbricated and folded series that includes Alpine-type Triassic and possibly early Liassic carbonates and flysch-facies deposits (Nalbantian Flysch). Porphyry and diabase dykes were intruded during the Trias. Middle Lias is unconformable on this deformed and intruded assemblage. The entire deformed package, making up the Tulcea Zone, is overridden from the south along a major thrust¹⁸ by the Macin Zone that was deformed largely during the Palaeozoic. Because the Baikalian (late Proterozoic) Moesian Platform with its gently deformed Phanerozoic cover is now

Fig. 1 The approximate area (vertically ruled) affected by mid-Mesozoic orogeny and thought to have coincided roughly with the present extent of the rocks representing the Cimmerian continent(s), the collision of which with Laurasia is believed to have caused the mid-Mesozoic orogeny. Localities for which stratigraphic, structural and petrologic data are given in the columns (1–15). The bold lines are sutures (dashed where inferred). Polarity information, where available, is indicated by black teeth on the overriding continent. Ages on the sutures are ages of ocean opening and closing for the segments limited by arrowheads and abbreviated as follows: Pz, Palaeozoic; T, Triassic; J, Jurassic; K, Cretaceous; P, Palaeocene; E, Eocene; O, Oligocene; M, Miocene; ↓, active subduction; e, early; m, medial; l, late. In the columns Pe denotes Permian. Horizontal zig-zag lines are unconformities.



located between North Dobrudja and the peri-Rhodopean Belt, relationships between these provinces are not clear.

Eastwards, along the strike from Dobrudja, the Triassic to Liassic Tauridian Flysch in the Crimean Mountains (Fig. 1, loc. 7; refs 19, 20) exhibits a strong, south-vergent fold and imbricated thrust structure²¹ unconformably overlain by Dogger sandstones. The same stratigraphy and structural picture is exhibited in the basement of the northern slope of the West Kubanian foredeep (Fig. 1, loc. 8; ref. 22), where a thick Middle–Upper Triassic flysch with volcanics has been folded during the Trias–Lias transition. Although similar lithologies of the same age have been encountered in deep drill-holes farther east (eastern shore of the Caspian Sea, Fig. 1, ref. 19), the structural picture there is not known. A strip is thus roughly defined north of the Black Sea characterised by deformed Triassic–Liassic flysch with volcanoclastic components and by diabase and porphyry intrusions.

Parallel with this strip, the southern shores of the Black Sea exhibit stronger, predominantly north-vergent mid-Mesozoic orogenic activity (at this time the Black Sea had not opened and its southern shores may have been much closer to the strip defined above, depending on the position of the Moesian Platform at the time). Around Balya in western Turkey (Fig. 1, loc. 4, ref. 23) Radelli mapped Permian limestones lying with thrust contact on Triassic and Lower Liassic rocks. The thrusts are truncated by an unconformity over which Middle Lias was deposited. Just east of Balya, north of Simav, the Eğrigöz granodiorite gives a Rb–Sr whole-rock age of 167 ± 14 Myr (ref. 24). In the Karagöl region, north of Ankara (Fig. 1, loc. 5; ref. 25), the same Permian limestones are overthrust onto an ophiolitic mélange containing spilites, serpentinites and blocks of the Permian limestones; this package, in turn, is seen in thrust contact on Hercynian basement and all are unconformably overlain by Lias. Other than the northeastern Pontide ophiolites of northeastern Turkey (see below), the suspected pre-Mesozoic ophiolite slivers of Mashad¹, northern Iran, and the serpen-

tinities along the Tanglha suture in Tibet (Fig. 1, loc. 15), this ophiolitic assemblage is the only occurrence I know that may represent Palaeo-Tethyan ocean floor material. Farther north-east, around the towns of Küre, Inebolu and Devrekani (Fig. 1, loc. 6; ref. 26), dioritic, granodioritic and gabbroic plutons have been mapped to intrude Liassic flysch. Pebbles of these intrusives are encountered in the basal conglomerate of transgressive Malm sequences. Dogger intrusions of 'granitic' composition are also known from the Greater Caucasus (Fig. 1, loc. 9; ref. 22). Blumenthal²⁷ has pointed out that the northernmost ophiolite outcrops in the northeastern Pontides in Turkey are of pre-Liassic age.

In central Iran there is abundant evidence for late Triassic–early Jurassic convergent plate–margin activity and Stöcklin has emphasised the importance of the 'cimmerian' movements in this sector of what he has designated the Central Domain, where (Fig. 1, loc. 10; refs 28, 29) late Triassic block-faulting is generally associated with thrusting and intense local folding. This deformation has often been cited as evidence for the rifting of central Iran from Arabia^{5,29}. However, the following observations indicate plate convergence: (1) the existence of strongly tectonised regions displaying Barrovian metamorphism (kyanite–biotite–garnet assemblages), accompanying deformation characterised by recumbent similar folding with a strong axial plane schistosity striking east–west in the southeastern extremity of the Sanandaj–Sirjan Zone (Fig. 1, loc. 11; refs 30, 31); (2) strong folding, thrusting and metamorphism of the basement of the Lut Block (central Iranian microcontinent of Takin³²) (Fig. 1, loc. 12; ref. 28). All of these events took place between the late Triassic and the early Jurassic^{28,30,31}. The coexistence of strong compressional deformation and the inferred block-faulting in central Iran during the late Triassic times may be explained in two different ways. First, the area of central Iran may have felt simultaneously the effects of the opening of the Bitlis/Zagros Ocean (a part of Neo-Tethys) and the closing of Palaeo-Tethys. A modern example of a similar association is

the Aegean region^{33,34}. Conversely, a basin-and-range regime may have existed in central Iran associated with complications of the convergent plate-margin regime related to the closing of Palaeo-Tethys. Similar complications are observed in the Neogene-Quaternary tectonics of the western US³⁵. Ultramafic rocks occurring in ?Carboniferous slates near Mashad, in northern Iran, were interpreted by Hsü¹ as possible representatives of the floor of Palaeo-Tethys.

Although the data are sparse and imprecise, in Afghanistan and Pakistan the Triassic-Jurassic transition is characterised by tectonism³⁶, a break in marine sedimentation³⁷ and intrusion of calc-alkaline igneous material such as Salang and Doab granites of western Hindu Kush Mountains (isotopic age: 210 ± 10 Myr; ref. 38). In western Hindu Kush, tectonism is characterised by a strong, north-vergent nappe tectonics of intra-Jurassic age³⁸ (Fig. 1, loc. 13). Farther east, in the Pamirs and the Karakorum (Fig. 1, loc. 14; refs 28, 39, 40) the late Triassic-early Jurassic orogenic activity is better documented: in the Pamirs the deformation was accompanied by the intrusion of calc-alkaline plutons and by a sharp facies change from shelf carbonates to a coarse clastic facies. In the Karakorum this latter facies is represented by the 'orogenic' Urdok Conglomerate which is the only coarse clastic sediment recognised in the Palaeozoic-early Mesozoic section⁴⁰.

Eastwards, the mid-Mesozoic zone of deformation enters the Tibetan Plateau, the geology of which is little known. Chang and Zeng⁴¹ indicate that there is a Mesozoic arc suture along the Tanglha Mountains (Fig. 1, loc. 15) with serpentinites and blueschists; isotopic ages along the Tanglha Mountains range from 210 to 107 Myr (ref. 41). Alternatively, Sengör and Kidd⁴² interpret the entire pre-Jurassic basement of Tibet as a giant accretionary mélange prism, first hit by a continent during the Himalayan collision. The Tanglha suture seems to continue along strike into the late Triassic Dien Bien Phu suture in southeastern Asia (Fig. 1; ref. 43). If the Tanglha suture represents a mid-Mesozoic continental collision (or arc-continent collision), we may continue the zone of mid-Mesozoic

orogeny into Tibet as a suture zone. Conversely, if the hypothesis of Sengör and Kidd⁴² is correct, then the Palaeo-Tethyan suture must have joined, from both sides, the large accretionary prism of Tibet, much as the present accretionary prism of Makran⁴⁴ links the Zagros and Indus-Tsangpo sutures. Palaeobiogeographic data (summarised in ref. 45) indicate that the welding of southern and northern Tibet occurred sometime between Permian and Jurassic.

Discussion and implications

The evidence reviewed above, when coupled with the marine magnetic and palaeogeographic arguments for the existence of Palaeo-Tethys, may be interpreted in terms of a late Triassic to medial Jurassic closure of this ocean. This closure occurred between Eurasia and the continental fragment(s) rifted from the northern side of Gondwanaland. Figure 2a shows the simplest possible interpretation of this continental fragment which I call here the Cimmerian continent after the Cimmerian Mountains that comprise the Macin Mountains, the Serpent Island and the Crimean Mountains. Clearly the Cimmerian continent may have had a very different geometry and have been composed of more than one fragment. The lack of precise simultaneity or of a regular age progression along the zone of mid-Mesozoic orogeny support the concept of a segmented Cimmerian continental archipelago and the apparent interruption in the zone of mid-Mesozoic deformation in northwestern Iran, the circum-central Iranian microcontinent suture (Fig. 1), and the evidence that Tibet may be characterised during this interval as a large accretionary prism indicate where such possible discontinuities may have been located. Neo-Tethys has also similarly closed by independent fragments of Gondwanaland. The pre-Alpine internal structure and possible width of the Cimmerian continent(s) are particularly difficult to assess because of the very strong Alpine overprint. In all areas where the mid-Mesozoic closure of Palaeo-Tethys caused deformation (Fig. 1) there has also been both intense Alpine orogenic deformation

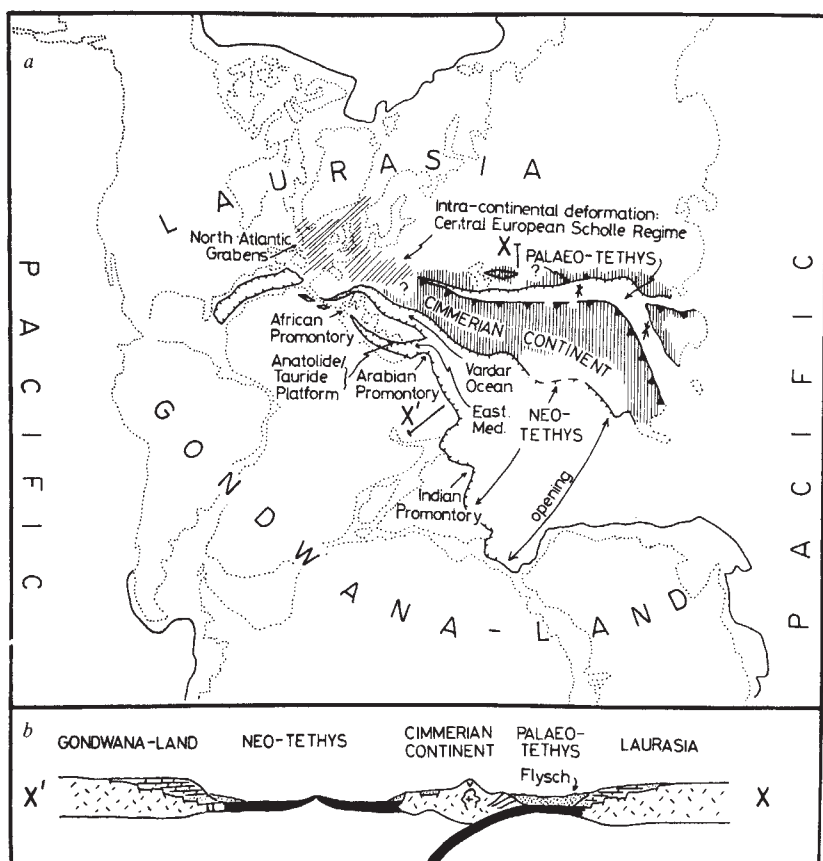


Fig. 2 a, Latest Triassic-early Jurassic reconstruction of Pangaea and Tethys (base map from ref. 54) showing the simplest possible interpretation of the internal geometry of the Tethyan realm. All the elements shown, particularly the poorly-defined Cimmerian continent, could have had different and probably more complex geometries. Vertically ruled area corresponds to that in Fig. 1, that is it represents the regions affected by mid-Mesozoic orogeny. Here the width of the Cimmerian continent is purely hypothetical and may be as much as three times too large. Obliquely ruled areas are regions of complex intracontinental extension inferred to have been partially related to the closure of Palaeo-Tethys. Lines with black triangles are subduction zones with the triangles on the upper plate. Lines with short hachures represent passive continental margins (no fit is intended between the southern and northern shores of Neo-Tethys due to lack of data). Dotted lines are the present continental shore-lines put in for reference. XX' indicates the location of the cross-section shown in b. b, A highly schematised cross-section showing the major elements considered in the text. Black denotes oceanic and short, irregular lines denote continental crust, +, Calc-alkaline intrusive; v, calc-alkaline volcanic rock.

and exceedingly complex post-collisional foreland deformation^{33,34,42,46,47}. These later events and the associated extensive volcanics and Neogene sedimentary cover seem to have obliterated or hidden the pre-Alpine characters of the Cimmerian continent(s) beyond any reasonable hope of coherent synthesis.

Because of the very fragmental and imprecise nature of the data, particularly from Iran eastwards, it is difficult to interpret the polarity(s) of Palaeo-Tethyan subduction zone(s). The location of relatively less deformed and largely unmetamorphosed flysch/molasse sequences to the north of the strongly deformed, metamorphosed and intruded, more 'internide-looking' predominantly north-vergent orogenic terrains as far east as Afghanistan may indicate that at least in this sector the Palaeo-Tethyan orogen was probably north-facing with a south-dipping subduction zone. In Tibet, if the regions north of the Tanglha Mountains are composed of late Palaeozoic-early Mesozoic accretionary mélange material, the subduction zone may have been dipping northward; but here data are so sparse that the question is open.

If the above proposed scheme is approximately correct, then it has interesting implications for two problems outside the Palaeo-Tethyan area. One is the complex, intracontinental zone of extension in extra-Alpine central and northwestern Europe (Fig. 1, loc. 1; refs 48–50). This broad zone of extension, characterised by a complicated pattern of block faulting (*Schollen*), became active during Liassic (?earlier) time and accelerated during the Malm. I suggest that this predominantly intracontinental extensional regime was possibly caused by the closure of wedge-shaped Palaeo-Tethys. If the pole of rotation of the Cimmerian continent(s) was located somewhere between North Dobrudja and Germany, as suggested by the observation that no early to mid-Mesozoic orogenic activity is known in Europe west of North Dobrudja, then the convergent plate boundary system, if not joined near the pole by other plate boundaries, must have become extensional once it crossed the pole. This extension should have increased away from the pole; indeed the early to mid-Mesozoic extension in the North Sea–North Atlantic regions was far more intense than it was in central Europe⁵⁰. The broad distribution of the associated strain may have been because extension was taking place in continental lithosphere and was following Permo–Triassic intracontinental grabens and shear zones⁵¹. It may, however, prove impossible to distinguish between extension caused by the initial

opening of the Atlantic and that related to the hinging of Palaeo-Tethys.

The second problem is related to the opening of Neo-Tethys. I suggest that if the Cimmerian continent or substantial segments thereof were sufficiently narrow, that is if they were essentially island arcs, the Triassic opening of Neo-Tethys may be interpreted as the opening of a series of back-arc basins or at least initiated by back-arc spreading processes (Fig. 2*b*). McKenzie and Weiss⁵² have pointed out that it seems particularly difficult to initiate constructive plate margins. They suggested that back-arc spreading may be the triggering mechanism for initiating major oceans, as back-arc spreading and seafloor spreading in major oceans seem to be the same process (see appendix in ref. 53). Although their main example, the South Atlantic, was unfortunate, the idea may be applicable, at least to parts of Neo-Tethys.

Conclusions

Palaeo-Tethys that must have existed during the Permo–Triassic between Laurasia and Gondwanaland closed during the late Triassic–mid-Jurassic possibly along one or more suture zones located within the zone of mid-Mesozoic orogeny shown in Fig. 1. As a result of very strong Alpine overprinting associated with the latest Mesozoic–Cenozoic demise of Neo-Tethys and because of the sparsity of data neither the precise geometry of the suture(s) nor that of the continent(s) that collided with Laurasia to close Palaeo-Tethys can be reconstructed. However, my interpretation permits a satisfactory explanation of a number of phenomena, notably the following: (1) The mid-Mesozoic convergent plate–margin activity in a relatively narrow band stretching from Rhodope and North Dobrudja to southeastern Asia and the observation that the Alpine–Himalayan sutures represent oceans opened not before the Triassic. (2) The mid-Mesozoic extensional regime in extra-Alpine central Europe which may be viewed as a partial consequence of the convergent plate–margin activity in the Palaeo-Tethyan realm. (3) The opening of Neo-Tethys which may be considered, at least in part, to be an example of back-arc spreading associated with the subduction of the Palaeo-Tethyan ocean floor.

I thank Professor Kenneth J. Hsü for stimulating my interest in the Palaeo-Tethys, Professors Kevin Burke and Clark Burchfiel for helpful comments.

Received 22 February; accepted 1 May 1979.

- Hsü, K. J. in *The Ocean Basins and Margins*, Vol. 4A (eds Nairn, A. E. M., Kanes, W. H. & Stehli, F. G.) 29 (Plenum, New York, 1977).
- Hsü, K. J., Nachev, I. K. & Vuchev, V. T. *Tectonophysics* **40**, 245 (1977).
- Bullard, E. C., Everett, J. E. & Smith, A. G. *Phil. Trans. R. Soc. A* **258**, 41 (1965).
- Smith, A. G. *Bull. geol. Soc. Am.* **82**, 2039 (1971).
- Dewey, J. F., Pitman, W. C., III, Ryan, W. B. F. & Bonnin, J. *Bull. geol. Soc. Am.* **84**, 3137 (1973).
- Smith, A. G. in *Implications of Continental Drift to the Earth Sciences*, Vol. 2 (eds Tarling, D. H. & Runcorn, S. K.) 977 (Academic, London, 1973).
- Smith, A. G. & Hallam, A. *Nature* **225**, 139 (1970).
- Norton, I. O. & Sclater, J. G. *J. geophys. Res.* (in the press).
- Boillot, G. & Capdevila, R. *Earth planet. Sci. Lett.* **35**, 151 (1977).
- Triimpy, R. *Am. J. Sci.* **275-A**, 209 (1975).
- Delaune-Mayer, M., Marcoux, J., Parrot, J.-F. & Poisson, A. in *Structural History of the Mediterranean Basins* (eds Bijou-Duval, B. & Montadert, L.) 79 (Technip, Paris, 1977).
- Marcoux, J. in *Int. Congr. on Sediment., Jerusalem* (in the press).
- Sengör, A. M. C., Dewey, J. F. & White, G. W. in *Rapports et Procès Verbaux des Reunions de CIESM* (in the press).
- Boulin, J. & Bouyx, E. *C. r. hébd. Seanc. Acad. Sci. Paris* **284**, 795 (1977).
- Argyriadis, I. *Bull. Soc. géol. Fr. Ser. 7*, 17, 56 (1975).
- Kockel, F., Mollat, H. & Walther, H. *Geol. Jb.* **89**, 529 (1971).
- Jacobshagen, V., Dürr, S., Kockel, F., Kopp, K.-O. & Kowalczyk, G. in *Alps, Apennines, Hellenides* (eds Closs, H., Roeder, D. & Schmidt, K.) 537 (Schweizerbart'sche, Stuttgart, 1978).
- Sandulescu, M. in *Geological Atlas of Alpine Europe and Adjoining Areas* (ed. Lemoine, M.) 427 (Elsevier, Amsterdam, 1978).
- Kotanski, Z. in *Geological Atlas of Alpine Europe and Adjoining Areas* (ed. Lemoine, M.) 545 (Elsevier, Amsterdam, 1978).
- Nalivkin, D. V. *Geology of the U.S.S.R.* **855** (Oliver & Boyd, Edinburgh, 1973).
- Brinkmann, R. *Stockwerkbau und Felderteilung* **14**, 477 (1973).
- Khain, V. E. *Am. J. Sci.* **275-A**, 131 (1975).
- Radelli, L. *Géol. Alp.* **46**, 169 (1970).
- Oztunali, Ö. *Inst. Univ. Fen. Fak. Monogr. (Tabii Ilimler Kismi)* **23**, 115 (1973).
- Fourquin, C. *Bull. Soc. géol. Fr. Ser. 7*, 17, 1058 (1975).
- M. T. A. *Explanatory Text of the Geological Map of Turkey*, 1/500,000; Sinop Sheet, 111 (MTA, Ankara, 1962).
- Blumenthal, M. M. T. A. *Mecmuası* **10**(1/33), 115 (1945).
- Stöcklin, J. in *Mém. h. sér. Soc. géol. Fr.* **8**, 333 (1977).
- Stöcklin, J. *Bull. Am. Ass. Petrol. Geol.* **52**, 1229 (1968).
- Vialon, P., Houchmand-Zade, A. & Sabzeji, M. *24th int. Geol. Congr.* **3**, 169 (1972).
- Ricou, L.-E., Braud, J. & Brunn, J. in *Mém. h. sér. Soc. géol. Fr.* **8**, 33 (1977).
- Takin, M. *Nature* **235**, 147 (1972).
- Dewey, J. F. & Sengör, A. M. C. *Bull. geol. Soc. Am. Part I*, **90**, 84 (1979).
- Sengör, A. M. C. *J. geol. Soc.* (in the press).
- Livaccari, R. F. *Geology* **7**, 75 (1979).
- Weipert, D., Wettekindt, H. P. & Wolfart, R. *Beih. Geol. Jb.* **92**, 99 (1970).
- Blaise, J., Bordet, P., Montenat, C., Desparmet, R. & Marin, P. *Mém. h. sér. Soc. géol. Fr.* **8**, 117 (1977).
- Boulin, J. & Bouyx, E. *Mém. h. sér. Soc. géol. Fr.* **8**, 87 (1977).
- Desio, A. *Mém. h. sér. Soc. géol. Fr.* **8**, 87 (1977).
- Desio, A. *Riv. Ital. Paleont.* **69**, 475 (1963).
- Chang, Chen-fa & Zheng, Shi-Lang *Sci. Sinica* **1**, 1 (1973).
- Sengör, A. M. C. & Kidd, W. S. F. *Tectonophysics* (in the press).
- Burret, C. F. *Earth planet. Sci. Lett.* **21**, 181 (1974).
- Farhoudi, G. & Karig, D. E. *Geology* **5**, 664 (1977).
- Ziegler, A. M., Scotese, C. R., McKerrrow, W. S., Johnson, M. E. & Bambach, R. K. *A. Rev. Earth planet. Sci.* **7**, 473 (1979).
- Molnar, P. & Tapponnier, P. *Science* **189**, 419 (1975).
- Dewey, J. F. *Tectonophysics* **40**, 53 (1977).
- Lotze, F. *Geologie Mitteleuropas*, 491 (E. Schweizerbart'sche Verlagsbuchhandlung, Stuttgart, 1974).
- Stille, H. in *3. Jahresber. Niedersächs. geol. Ver.* (Hannover) 141 (1910).
- Whiteman, A. J., Rees, G., Naylor, D. & Pegrum, R. M. *Norges geol. undersøk.* **136**, 137 (1975).
- Arthaud, F. & Matte, P. *Bull. geol. Soc. Am.* **88**, 1305 (1977).
- McKenzie, D. P. & Weiss, N. O. *Geophys. J. R. astr. Soc.* **42**, 131 (1975).
- Molnar, P. & Atwater, T. *Earth planet. Sci. Lett.* **41**, 330 (1978).
- Smith, A. G., Briden, J. C. & Drewry, G. E. *Organisms and Continents Through Time*, 1 (Palaeontological Association, London, 1973).

Sand on the southern Mediterranean Ridge: proximal basement and distal African–Nile provenance

Daniel Jean Stanley & Harrison Sheng

Division of Sedimentology, Smithsonian Institution, Washington, D.C. 20560

Mahmoud M. Kholief

Egyptian Petroleum Research Institute, Cairo, Egypt

Petrographic analysis of Quaternary terrigenous sand layers in eastern Mediterranean cores reveals distinct mineralogical differences between the Egyptian Shelf–Nile Cone region and the southern part of the Mediterranean Ridge. A compositionally and texturally immature suite in Ridge cores, mixed with a Nile-derived assemblage, identifies a fresh non-recycled mineral component derived from proximal igneous and metamorphic surface or near-surface exposures, probably in the south-central Ridge area rather than from distal African sources. The presence of such basement terrains would be consistent with a compressive thrust-belt origin for this part of the Mediterranean Ridge.

THE Mediterranean Ridge, the broad arcuate swell in the eastern Mediterranean extending from the Ionian Sea to the eastern Levantine Basin, constitutes a distinct topographic and tectono-stratigraphic province between the seismically active southern European and more stable North African margins. The characteristic hummocky ridge-and-swell topography of the Ridge is the surface expression of what appears in seismic profiles as a folded and faulted sediment complex^{1,2}. The contact between deformed sedimentary units of the southern part of the Ridge and the more gently stratified, Plio–Quaternary series forming the western Nile Cone tends to be well-defined³ and delineates the northwestern margin of the Herodotus Basin plain (Fig. 1). A mid-oceanic ridge origin⁴ can be ruled out for the Mediterranean Ridge, but there is considerable diversity of opinion as to the structural and stratigraphic configuration^{5–7}. The various hypotheses pertaining to the geological development of this feature are reviewed by Finetti⁸ and Stride *et al.*⁹.

All the schemes invoke important large-scale tectonic displacement in the late Tertiary to the present, and one would expect that evidence of these geologically recent events should be recorded in sedimentary sequences deposited in this area. This article focuses on the provenance of terrigenous sand layers of Pleistocene age on the southern part of the Ridge at the approximate midpoint between the more obvious potential sources of sediment: Nile Delta, Crete, Cyprus and the Eratosthenes Seamount. This sector has more specifically been termed the Hellenic Outer Ridge⁹. Pleistocene and possibly Upper Pliocene core sections recovered in this part of the Ridge in 1970 at DSDP Site 130 (Fig. 1) have been interpreted as distal Nile Cone deposits which were uplifted <1 Myr ago and subsequently deformed². Arguments in support of this scheme are based partly on regional correlation of sub-bottom reflectors of probable Messinian and post-Miocene age, but more specifically on mineralogical comparisons between DSDP Sites 130 and 131 at the base of the Nile Cone (ref. 2, p 735 and 1033;

ref. 10) and on the pollen content in Pleistocene samples at Site 130¹¹. These studies all emphasise the North African, largely Nile, affinity of the southern Ridge deposits. The difference in relative percentages of the dominant heavy minerals (primarily an amphibole–pyroxene–epidote–opaque mineral suite) between Nile Cone and Ridge samples is attributed by Ryan and others² to differences in grain size and to changes of source terrain drained by the Nile River during the Quaternary.

This study considers heavy and light minerals from terrigenous sand layers, mostly turbidites, in DSDP Site 130 and 131 cores and in Pillsbury 6508 and 6510 and Vema 10 piston and gravity cores on the outer Egyptian Shelf, Nile Cone and Herodotus Basin plain, and shallower marine palimpsest deposits in Chain 119 grab samples on the inner to middle Egyptian Shelf in front of the Nile Delta (Fig. 1). Observed regional differences indicate a more complicated provenance-dispersal of Pleistocene sediment on the southern part of the Mediterranean Ridge than previously envisioned.

Petrographic observations

Most Pleistocene sand layers in the Nile Cone and the Mediterranean Ridge core sections are of bioclastic origin¹², and our investigation considers only terrigenous or mixed terrigenous–bioclastic layers. Of about 225 sand layers sampled in the cores shown in Fig. 1, only 47 examined had a sufficiently heavy mineral content for grain count (Table 1). Data are based on the calculated 62 to 250 μ m heavy mineral fraction of each layer to minimise size-sorting effects. Heavy fractions were separated by the standard heavy liquid method and 300 grains were identified. The listed data shows that the study area occupies a province dominated by pyroxenes, amphiboles, epidotes and opaque minerals; these four account for more than 90% of the total heavy minerals in most samples, as was indicated in earlier investigations^{2,10,13,14}.

Vertical changes were observed in some cores (Table 1): for example, in DSDP Site 131 from the lower Nile Cone, there is a slight downcore increase in pyroxenes and epidotes and an increase in amphiboles; in Pillsbury core 6510–18 on the outer Egyptian Shelf, there is a slight downcore increase in pyroxenes and amphiboles, and a more random change in epidote content. Regional–spatial changes are more apparent and include a general decrease in the relative percentage of opaque minerals between the outer edge of the shelf and distal Nile Cone stations. Overall, the Nile Cone samples in grabs, short piston and gravity cores and in DSDP Site 131 are genetically related to the same general assemblage as the Nile River valley, Nile Delta and coastal sequences^{13–17}.

The most pronounced changes in heavy mineral assemblages occur between the inner and outer Egyptian Shelf and between the base of the Nile Cone and the southern part of the Mediterranean Ridge. These regional distinctions become evident when considering the ranges of relative percentages of minerals in the

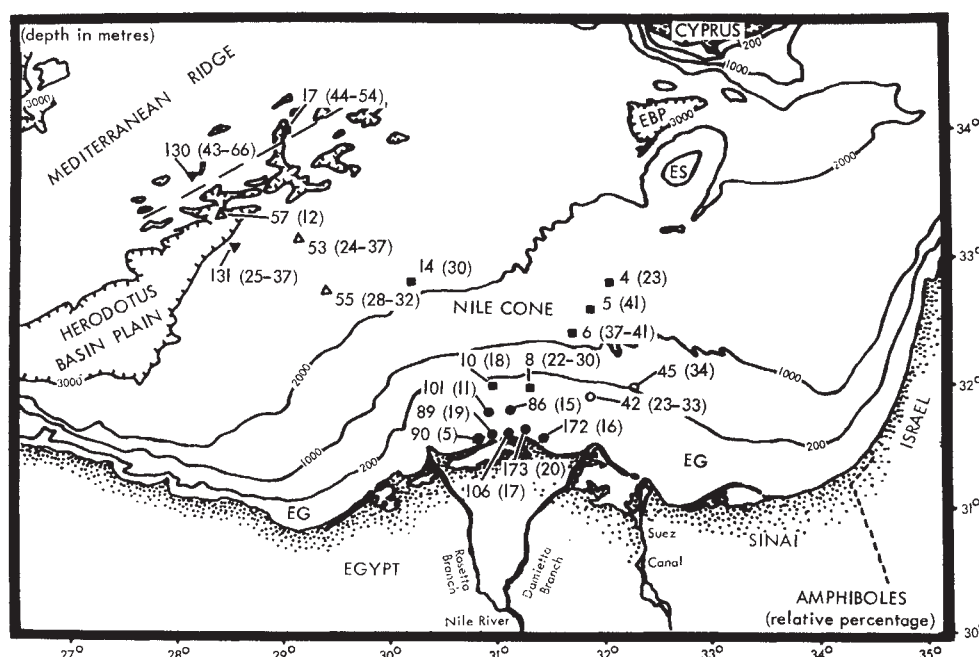


Fig. 1 Chart of the southeastern Mediterranean showing sample station locations on the Egyptian Shelf (EG), Nile Cone and southern part of the Mediterranean Ridge. ES, Eratosthenes Seamount; EBP, Eratosthenes Basin plain. Numbers refer to core and grab sample station and, in parentheses, the relative percentage of amphiboles in the non-opaque heavy mineral assemblage (data from Table 1). Note increased amphibole content in Ridge samples. ●, Chain 119; ○, Pillsbury 6508; ■, Pillsbury 6510; △, Vema 10; ▼, DSDP Leg 13.

three major topographic sectors of the study area (Table 1):

(A) On the Egyptian Shelf (10 stations): clinopyroxenes (8.5–62.8), orthopyroxenes (0.0–7.9), amphiboles (5.0–33.3), epidotes (6.5–37.0), garnets (0.0–46.3), zircon–toulmaline–rutile, or ZTR (0.0–6.5), apatite (0.0–1.6), opaques (23.0–70.0).

(B) On the Nile Cone proper (nine stations): clinopyroxenes (12.9–46.5); orthopyroxenes (0.0–3.3); amphiboles (11.9–40.8); epidotes (4.8–30.6); garnets (0.0–2.5); ZTR (0.0–4.7); apatite (0.0–1.6); opaques (11.0–60.0).

(C) On the southern Mediterranean Ridge (two stations): clinopyroxenes (10.0–17.3); orthopyroxenes (0.4–2.0); amphiboles (43.4–65.8); epidotes (8.5–26.5); garnets (0.4–3.6); ZTR (0.4–3.6); apatite (1.6–2.2); opaques (47.0–68.0).

Nile Cone (B) samples display a narrower range in relative percentage and smaller core-to-core differences (particularly clino- and orthopyroxenes and garnets) than the Egyptian Shelf (A) samples. The southern Ridge (C) samples (DSDP Site 130 and Pillsbury 6510–17) are distinguished from Nile Cone–Herodotus Basin plain (B) samples by enhanced proportions of amphiboles and lower percentages of pyroxenes as noted earlier by Ryan *et al.*² and Bartolini *et al.*¹⁰. Moreover, a careful evaluation of specific heavy mineral types in our Ridge core sample separations reveals other significant differences: a somewhat higher proportion of brown hornblende, serrated or saw-tooth pyroxenes (Fig. 2a), angular apatite (Fig. 2c); the presence of a few glaucophane grains; and a markedly higher proportion of both garnet and opaque minerals in the DSDP Site 130 samples. Examples of the more typical rounded pyroxene and apatite grains in the Nile Cone are shown for contrast in Fig. 2b and d.

The compositional differences between sands in the lower Nile Cone–Herodotus Basin plain sector and those on the southern part of the Ridge also are recorded in the light mineral fractions. There is a relatively higher proportion of fresh to slightly altered and angular feldspar grains on the Ridge (Fig. 2e) than in DSDP Site 131 and other Nile Cone core samples where subangular to rounded altered grains prevail (Fig. 2f). In DSDP Site 131 and other Nile Cone core samples, there is a strong wavy extinction: normal–weakly wavy extinction ratio of about 2:1. Two samples of DSDP Site 130 contain quartz with primarily normal or weakly wavy extinction, and three with strongly wavy extinction; most quartz in Ridge core P6510–17 also displays strong wavy extinction.

More distinct regional differences are recorded by the nature and frequency of quartz inclusions: samples from DSDP Site

130 and P-6510–17 tend to show fewer solid (rutile needles, zircon, apatite) and fluid–gaseous inclusions than in DSDP Site 131 and Nile Cone and outer Shelf samples (Fig. 3a, b) and in comparable Nile Delta sands as earlier detailed by Kholief¹⁸. Some quartz grains from Ridge core P6510–17 (sands sampled at 149 and 452 cm from top of core) and DSDP Site 130 (sample 4–2) are characterised by angular overgrowths over euhedral and subhedral crystals (Fig. 3c), unlike some quartz from the

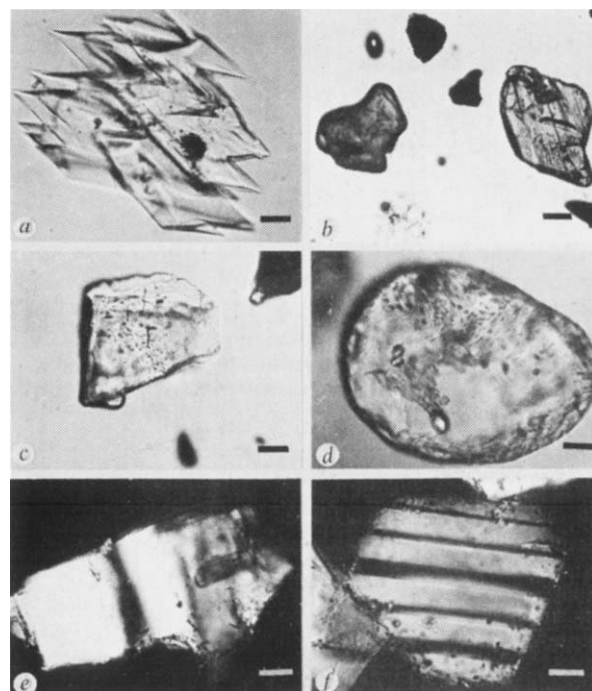


Fig. 2 Photomicrographs showing details of some heavy and light minerals from the southern Mediterranean Ridge (a, c, e) and Nile Cone (b, d, f) core samples. a, Saw-tooth amphibole (Pillsbury 6510–17, 452 cm) scale bar, 10 µm; b, subrounded amphibole (DSDP Site 131–1–6, 129–131 cm) scale bar, 100 µm; c, angular apatite (DSDP Sites 130–5–3, 125–126 cm) scale bar, 25 µm; d, rounded apatite (DSDP Site 131A–4–2, 117–119 cm) scale bar, 15 µm; e, fresh, angular plagioclase feldspar (Pillsbury 6510–17, 452 cm) scale bar, 15 µm; f, subrounded, worn-edged plagioclase feldspar (Vema 10–57, 690–691 cm) scale bar, 15 µm.

Table 1 Heavy mineral data from samples collected on the southern Mediterranean Ridge, Nile Cone and Egyptian Shelf

Region	Core no.*	Depth† (cm)	Relative percentage of heavy non-opaque minerals									% Opaques	% Non-opaques
			Orthopyroxenes	Clinopyroxenes	Amphiboles	Epidote	Garnet	ZTR‡	Sphene	Apatite	Others§		
Southern sector of Ridge	13-130-4-2	96-97	0.9	17.3	65.8	8.5	0.9	0.4	—	1.7 ^a	4.5	47.0	53.0
	13-130-5-2	90-91	1.8	16.4 ^c	43.4	26.5	1.3	3.1	0.4	1.8 ^a	5.3	68.0	32.0
	13-130-5-3	125-126	0.8	13.5 ^c	59.8 ^b	10.4	3.6	3.6	—	1.6 ^a	6.7 ^d	51.0	49.0
	P6510-P17	149	2.0	10.0 ^c	53.8 ^b	21.3	0.4	1.6	0.8	—	10.9 ^d	8.0	92.0
	P6510-P17	452	0.4	14.5 ^c	44.1 ^b	21.1	0.9	2.2	—	2.2 ^a	14.6 ^d	28.0	72.0
Western Nile Cone-Herodotus Basin Plain	13-131-1-6	18-20	0.4	46.5	25.2	18.4	2.0	2.0	—	1.6	3.9	41.0	59.0
	13-131-1-6	129-131	0.4	41.3	25.4	24.1	2.5	4.7	—	0.4	1.2	42.0	58.0
	13-131A-1-2	55-57	0.3	45.6	32.9	16.1	0.3	0.3	1.4	0.3	2.8	32.0	68.0
	13-131A-1-2	81-83	0.4	37.2	32.1	25.8	2.0	—	1.2	—	1.3	24.0	76.0
	13-131A-1-2	122-124	0.4	43.1	36.5	15.4	0.4	0.4	—	0.4	3.4	28.0	72.0
	13-131A-1-2	145-147	—	43.8	30.5	15.4	1.3	1.3	0.8	—	6.9	25.0	75.0
	13-131A-4-2	8-10	—	43.9	29.2	18.3	1.4	1.0	0.3	0.3	5.6	34.0	66.0
	13-131A-4-2	117-119	—	45.4	24.9	26.5	1.8	1.1	—	—	0.3	26.0	74.0
	13-131A-4-2	137-139	—	43.7	34.1	18.5	0.3	0.7	—	—	2.7	21.0	79.0
	V10-53	330-331	2.5	30.2	23.5	16.9	0.4	2.1	—	—	24.4	28.0	72.0
	V10-53	372-373	—	12.9	33.2	16.1	0.9	—	—	0.9	36.0	17.0	83.0
	V10-53	384-385	2.5	26.2	26.9	23.3	—	1.1	0.4	0.4	19.2	24.0	76.0
	V10-53	645-646	0.9	17.9	36.8	16.1	0.9	1.3	—	—	26.1	11.0	89.0
	V10-55	240-241	3.3	25.4	28.2	16.7	0.5	2.9	0.5	—	22.5	30.0	70.0
	V10-55	281-282	—	20.3	29.3	30.6	—	0.4	—	0.4	19.0	20.0	80.0
	V10-55	410-411	0.4	24.3 ^c	31.7	28.4	0.8	0.8	—	0.4	13.2	27.0	73.0
	V10-57	690-691	2.7	35.2	11.9	20.1	—	0.9	—	—	29.2	25.0	75.0
	P6510-P14	1128-1129	2.4	26.4	29.5	16.9	0.8	2.4	0.4	0.4	20.8	22.0	78.0
Eastern Nile Cone	P6508-45	10	—	42.8	34.2	22.8	—	—	—	—	0.2	49.0	51.0
	P6510-P4	582	1.4	22.5	23.0	4.8	—	0.5	0.5	—	47.3	14.0	86.0
	P6510-5	130-131	1.2	40.7	40.6	9.8	—	—	—	1.2	6.5	60.0	40.0
	P6510-6	122-123	—	45.7	37.0	14.4	0.9	1.3	—	—	0.7	30.0	70.0
	P6510-6	123-124	1.8	46.5	37.2	8.2	—	—	—	0.4	5.9	42.0	58.0
	P6510-6	142	—	30.2	40.8	20.1	—	—	0.4	—	8.5	37.0	63.0
Outer Egyptian Shelf	P6510-8	1-2	0.4	50.0	24.5	16.5	2.9	2.9	0.8	—	2.0	41.0	59.0
	P6510-8	5-7	—	48.0	25.0	18.4	3.5	3.4	0.4	—	1.3	38.0	62.0
	P6510-8	84-85	0.4	33.3	22.1	37.0	3.2	0.4	—	0.9	2.7	26.0	74.0
	P6510-8	90-91	—	38.8	24.2	31.1	1.6	2.4	—	1.6	0.3	32.0	68.0
	P6510-8	99-100	—	42.1	24.5	23.1	1.1	1.0	0.7	0.3	7.2	34.0	66.0
	P6510-8	109-111	—	37.7	26.9	26.9	1.2	0.4	0.4	0.4	6.1	38.0	62.0
	P6510-8	112-113	—	45.2	30.3	18.1	1.0	0.7	—	—	4.7	23.0	77.0
	P6510-10	90	—	62.8	18.0	6.5	3.7	—	—	0.4	8.6	37.0	63.0
	P6508-42	2	—	45.1	22.9	20.1	3.8	3.8	—	—	4.3	64.0	36.0
	P6508-42	70	—	39.3	33.3	19.6	2.2	4.4	—	—	1.2	70.0	30.0
	P6508-42	100	—	43.1	26.1	26.6	2.7	0.8	0.4	—	0.3	37.0	63.0
Inter-Mid Egyptian Shelf	CHN 119-II, sta 86 Grab 11		0.5	18.7	14.9	28.8	13.0	6.2	1.0	—	16.9	46.0	54.0
	CHN 119-II, sta 89 Grab 13		4.2	12.7	19.3	33.5	0.9	0.9	0.5	—	28.0	29.0	71.0
	CHN 119-II, sta 90 UW 24		—	8.5	5.0	12.4	46.3	6.5	—	—	21.3	52.0	48.0
	CHN 119-II, sta 101 UW 34		1.5	25.1	10.8	18.7	10.8	4.9	—	—	28.2	42.0	58.0
	CHN 119-II, sta 106 Grab 15		7.9	26.5	16.7	22.3	—	2.3	—	0.5	23.8	31.0	69.0
	CHN 119-II, sta 172 Grab 35		1.8	41.3	15.5	20.0	0.4	0.9	—	—	20.1	36.0	64.0
	CHN 119-II, sta 173 UW 44		2.6	39.3	19.6	20.5	1.3	1.7	1.3	—	13.7	31.0	69.0

* Key to sample stations: 13 refers to DSDP 1970 leg 13 Sites 130, 131 and 131A, with drill core and section numbers as in ref. 2; P, University of Miami cruise Pillsbury 6508 and 6510 piston and gravity cores²⁰; V, Lamont-Doherty Geological Observatory cruise Vema 10 piston cores¹²; CHN, Woods Hole Oceanographic Institution cruise Chain 119 grab samples¹⁹.

† Depth in centimetres from piston and gravity core tops or DSDP core section top.

‡ Zircon + tourmaline + rutile.

§ Includes altered as well as unidentified grains.

^a Angular apatite; ^b brown hornblende present to common; ^c angular and saw-tooth pyroxenes; and ^d presence of few glaucophane grains.

Nile Cone and Egyptian Shelf which display partial rounded overgrowths on rounded grains (Fig. 3d). Some angular quartz grains of the Ridge show no evidence of wear. Also significant are angular and euhedral quartz and feldspar grains (found only in DSDP Site 130, samples 4-2 and 5-3) that display a type of radiating crack pattern with fissures that extend from the centre towards the edge of the grain; the intensity of fissuring decreases outward (Fig. 3e). This pattern is in marked contrast with quartz, usually rounded, in the Delta and on the Shelf and Nile Cone, which more commonly displays one or two generations of cracks (Fig. 3f).

Evidence for mixed Nile-proximal provenance

The mineralogical composition of terrigenous sand-sized material in surficial Egyptian Shelf and Quaternary Nile Cone samples, including DSDP Site 131, are genetically related to a Nile apportion. The large station-to-station differences in the relative percentage of heavy minerals in the surficial Egyptian Shelf samples primarily reflect size sorting and density concentration (lag) effects due to sea floor processes that have reworked both modern and relict sediments¹⁹. During the Pleistocene, Nile-

derived sandy fluvial, deltaic and eolian deposits were emplaced downslope by gravity-induced resedimentation processes^{2,10,12,20}. The observed fluctuations in relative percentages of heavy minerals between samples within any core and between piston core localities on the different parts of the Nile Cone proper more closely record differences of processes than changes in provenance. This is demonstrated, for example, by the progressive diminution downslope, in the same textural grade, of denser opaque minerals. Both heavy and light mineral suites in Nile Cone Quaternary short piston core and DSDP Site 131 samples are essentially comparable with those described in the main Nile by Shukri^{15,16}, and in the sediments of the Delta and related Nile-derived older sections^{14,17,21}, which show a contribution (including relatively high percentages of pyroxenes) from the Blue Nile, Atbara and other tributaries draining the Abyssinian volcanic source terrains.

Although composition of the probable lower to middle Pleistocene sands at DSDP Site 130 and Pleistocene sands in Pillsbury core 6510-17 is broadly similar to the Nile-derived sediment on the Nile Cone, our investigation emphasises some significant mineralogical differences on the southern part of the Ridge. We believe that the distinct suite of mineral species in the Ridge samples indicates input from a more proximal source terrain in addition to the Nile-derived sands. The much higher proportion of amphiboles in southern Ridge cores is not, as previously suggested, due only to a finer grain size in this more distal sector; examination of comparable textural grades in Nile Cone DSDP Site 131 and Vema 10-53, 55 and 57 core samples all show substantially lower proportions of amphiboles than in DSDP Site 130 and P6510-17 core samples (Fig. 1). Furthermore, the greater proportion of angular grains of both heavy and light components (apatite, serrated pyroxenes, quartz, feldspar), an increased percentage of unaltered feldspar grains, apatite and brown hornblende, the presence of glaucophane grains, angular quartz overgrowth on euhedral and subhedral quartz grains, a radiating crack pattern within some euhedral quartz and feldspar grains, and a generally lower proportion of solid inclusions in quartz grains identify a mineralogical assemblage distinct from that in the modern to pre-Quaternary main Nile, Nile Delta and Nile coastal deposits.

The angular, serrated and non-altered state of various mineral species in the Ridge samples records minimum weathering-abrasion effects; this is in contrast with the reworked, subangular to rounded and often altered grains on the Egyptian Shelf and Nile Delta that invariably show considerable effect of one or multi-cycle fluvial or eolian (or both) transport. The immature mineral fraction in DSDP Site 130 and P6510-17 sands records the introduction of material from more proximal plutonic igneous and possibly also metamorphic sources rather than from distal African terrains. The compositional mix of reworked Nile-derived material and 'fresh' components would imply the addition of new sediment at some point between the Nile Delta-Egyptian Shelf sector and the Ridge core localities. As most of the sands examined in cores DSDP Site 130 and P6510-17 were emplaced by turbidity currents and associated gravity induced mass-flow processes, we are required to consider some modification of the direct downslope SE to NW dispersal path that has prevailed in this sector during the late Quaternary^{12,20}.

Two-phase dispersal pattern

Where did the fresh sedimentary material enter the transport system? An extensive seismic survey of the Nile Cone by Ross and Uchupi³ does not reveal basement series at or near the surface between the Nile Delta and the Ridge, and we exclude a southerly position for the introduction of fresh minerals. Transport from a northern Hellenic Arc-Crete source is precluded on the basis of the heavy mineral composition which is significantly different in that sector, a view also adopted by Bartolini *et al.*¹⁰. Cyprus and the Eratosthenes Seamount, located approximately 300 km NE and E of the sample sites, while not excluded, seem to be unlikely source areas. We can do

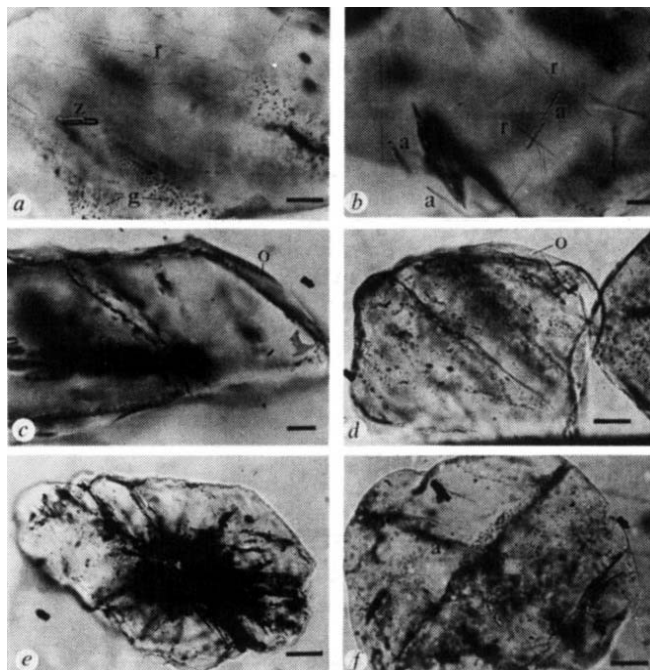


Fig. 3 Photomicrographs showing details of quartz from Nile Cone and Mediterranean Ridge core samples. *a*, Quartz with inclusions of partially altered zircon (*z*), segmented rutile needles (*r*) and orientated minute gas (*g*) bubbles (outer Egyptian Shelf, Pillsbury 6510-10, 90 cm); rutile segments are of variable length and arranged in a regular pattern; scale bar, 55 μ m. *b*, Quartz with non-orientated inclusions of apatite (*a*) and rutile (*r*) needles (outer Egyptian Shelf, Pillsbury 6510-10, 90 cm) scale bar, 20 μ m. *c*, Angular secondary overgrowth (*O*) on subhedral quartz, with orientated cracks and green solid inclusions (southern Mediterranean Ridge, Pillsbury 6510-17, 452 cm) scale bar, 10 μ m. *d*, Rounded secondary overgrowth (*O*) partially surrounding a subrounded quartz grain, which shows a system of orientated cracks with secondary inclusions (outer Egyptian Shelf, Pillsbury 6510-10, 90 cm) scale bar, 55 μ m. *e*, Subhedral quartz with radiating crack pattern displaying fissures that extend outwards from the centre of the grain (southern Mediterranean Ridge, DSDP Site 130-5-2, 90-91 cm) scale bar, 50 μ m. *f*, Round quartz grain displaying two generations of cracks lined with secondary inclusions, and an apatite (*a*) solid inclusion (outer Egyptian Shelf, Pillsbury 6510-10, 90 cm) scale bar, 55 μ m.

little more, at this point, than invoke a two-phase dispersal pattern: transport downslope of recycled African sand on the Nile Cone surface in a direction away from the Nile Delta to an area north of the present Herodotus Basin plain during a first stage of Nile Cone progradation², and then a remobilisation and subsequent short-distance transport of this Nile-derived material to which a distinct igneous and metamorphic fraction has been added in the Ridge area. The topographically complex configuration of the Ridge is of geologically recent origin, and it is possible that some resedimentation occurred during the Pleistocene in the area where transport would not be possible on the present Ridge surface.

Immature mineral assemblages in Pleistocene sands of the short P-6510-17 core, about 90 km NE of the much older Pleistocene DSDP Site 130 locality, strengthens the hypothesis of yet undefined igneous-metamorphic sources in the south-central Ridge area. The possibility of a surface or near-surface exposure of sub-sedimentary basement on the Ridge should not be readily dismissed although no specific evidence for this has been provided by gravity measurements or seismic profiles. Our findings, not inconsistent with compressive thrust-belt⁸ and décollement²² reconstructions, suggest that possibility of some movement of basement, perhaps as part of an allochthonous

slice. The petrographic evidence indicating local input should be considered in developing an interpretive scheme for the Mediterranean Ridge.

We thank the NSF (DSDP Site 130 and 131 cores), University of Miami School of Marine and Atmospheric Science (P-6510 cores), Lamont-Doherty Geological Observatory of Columbia University (Vema 10 cores) and Woods Hole Oceanographic Institution (Chain 119 grab samples) for generously providing samples. Dr Andrés Maldonado assisted in sampling and Drs W. G. Melson and D. A. Ross reviewed the manuscript. This study, part of the Mediterranean Basin (MEDIBA) Project, was funded by a Smithsonian Research Award (FY-79 grant 89191028).

Received 26 February; accepted 4 May 1979.

1. Ryan, W. B. F. *et al.* in *The Sea* Vol. 4, 387-492 (ed. A. E. Maxwell) (Wiley-Interscience, New York, 1971).
2. Ryan, W. B. F. *et al.* *Init. Rep. DSDP Leg 13*, 1-1447 (1973).

3. Ross, D. R. & Uchupi, E. *Am. Ass. Petrol. Geol. Bull.* **61**, 872-902 (1977).
4. Emery, K. O., Heezen, B. C. & Allan, T. D. *Deep-Sea Res.* **13**, 173-192 (1966).
5. Rabinowitz, P. D. & Ryan, W. B. F. *Tectonophysics* **10**, 585-608 (1966).
6. Mulder, C. J. in *Messinian Events in the Mediterranean* (ed. Drooger, C. W.) 44-59 (North-Holland, Amsterdam, 1973).
7. Sancho, J. *et al.* *Earth planet. Sci. Lett.* **18**, 189-204 (1973).
8. Finetti, I. *Boll. Geogr. Teor. Appl.* **18**, 31-65 (1976).
9. Stride, A. H., Belderson, R. H. & Kenyon, N. J. *Phil. Trans. R. Soc. A* **284**, 255-285 (1977).
10. Bartolini, C. *et al.* *Sedimentology* **22**, 205-236 (1975).
11. Rossignol-Strick, M. *Sci. Geol. Bull.* **27**, 135-145 (1974).
12. Maldonado, A. & Stanley, D. J. in *Sedimentation in Submarine Canyons, Fans and Trenches* (eds Stanley, D. J. & Kelling, G.) 239-157 (Dowden, Hutchinson & Ross, Stroudsburg, 1978).
13. Hilmy, M. E. *J. Sed. Petrol.* **21**, 109-129 (1951).
14. Kholief, M. M., Hilmy, E. & Shahat, A. *J. Sed. Petrol.* **39**, 1520-1529 (1969).
15. Shukri, N. M. *Q. J. geol. Soc. Lond.* **106**, 511 (1950).
16. Shukri, N. M. *Bull. Inst. Fouad Ier Desert* **1**, 39-67 (1951).
17. Chumakov, I. S. *Trans. Acad. Sci. U.S.S.R.* **170**, 1-113 (1967).
18. Kholief, M. M. *N. Jb. Mineral Mh.* **9**, 418-424 (1974).
19. Summerhayes, C. P. *et al.* *Mar. Geol.* **27**, 43-65 (1978).
20. Stanley, D. J. & Maldonado, A. *Nature* **266**, 129-135 (1977).
21. Kholief, M. M. in *The Mediterranean Sea—A Natural Sedimentation Laboratory* (ed. Stanley, D. J.) 345-354 (Dowden, Hutchinson & Ross, Stroudsburg, 1972).
22. Kenyon, N. H. & Belderson, R. H. in *Structural History of the Mediterranean Basins* (eds. Bijou-Duval, B. & Montadert, L.) 233-240 (Technip, Paris, 1977).

Characterisation of deletions which affect the expression of fetal globin genes in man

Edward F. Fritsch, Richard M. Lawn & Tom Maniatis

Division of Biology, California Institute of Technology, Pasadena, California 91125

Deletions in the DNA of individuals with hereditary persistence of fetal haemoglobin (HPFH) and $\delta\beta$ -thalassaemia have been mapped as a means of identifying regulatory sequences involved in the switch from fetal to adult globin gene expression. The end points of these deletions have been precisely located with respect to restriction endonuclease cleavage sites within and surrounding the γ -, δ - and β -globin genes in normal human DNA and the deletion maps were used to obtain definitive evidence for the physical linkage of the fetal and adult β -like globin genes in the order 5'^G γ - γ - δ - β 3'. Correlation of haematological data and the location of deletions in two cases of HPFH and one case of $\delta\beta$ -thalassaemia suggest that a region of DNA located near the 5'-end of the δ -globin gene may be involved in the suppression in cis of γ -globin gene expression in adults. The interpretation of a second case of $\delta\beta$ -thalassaemia is complicated by the fact that the deletion removes the γ -gene in addition to the region near the 5'-end of the δ -globin gene.

HUMAN globin genes provide a unique system for studying the nature of mutations which alter the normal programme of differential gene expression during mammalian development. For example, in normal individuals the fetal β -like globin genes, designated γ and γ , are turned down at 32-34 weeks of gestation, while the adult β -like genes, designated δ and β , are turned up. Certain genetic disorders have been described in which this control 'switch' is altered (see refs 1, 2 for reviews); in that designated hereditary persistence of fetal haemoglobin (HPFH)¹⁻¹², often neither the δ - nor the β -globin chains are synthesised in adult erythroid cells, but one or both of the

γ -globin genes continue to be expressed, compensating for the β -globin deficiency. In $\delta\beta$ -thalassaemia, no δ - or β -globin chain synthesis is detected in adults, but in contrast to HPFH, compensation by continued fetal globin synthesis is incomplete. The resulting imbalance of α - and non- α -globin polypeptide synthesis in homozygous $\delta\beta$ -thalassaemia leads to anaemia and red cell abnormalities characteristic of thalassaemia syndromes¹. Differences in the pattern of globin gene expression in individuals with HPFH and $\delta\beta$ -thalassaemia led to the proposal that sequences necessary for the repression of γ -globin gene expression in the adult might be deleted in HPFH DNA⁷. In fact, when DNAs from individuals homozygous for HPFH or $\delta\beta$ -thalassaemia were examined by solution hybridisation experiments, extensive deletions of β -globin gene sequences were observed³⁻⁶. This result was confirmed^{9,10} using the blot hybridisation technique of Southern¹³ to map single copy genes^{14,15}. Mapping deletions in HPFH and $\delta\beta$ -thalassaemia may provide insights into the mechanism of the switch from fetal to adult globin gene expression.

Advances in physical mapping and gene cloning procedures have made it possible to determine the linkage arrangement of human fetal and adult globin genes and have provided hybridisation probes for examining DNA which lies outside gene sequences. Previous genetic evidence^{1,2} and structural analysis of the fusion proteins Lepore¹⁶ and Kenya¹⁷ suggested that the fetal and adult β -like globin genes are linked. Physical linkage of the δ - and β -globin genes^{18,19} and of the γ - and γ -globin genes²⁰ was demonstrated using blot hybridisation¹³ and gene cloning²¹ procedures. Physical linkage between δ - β and γ - γ gene loci has not been demonstrated.

In this report we use cloned complementary DNA and genomic DNA fragments as probes in blot hybridisation experiments to map the end points of deletions in DNA from individuals homozygous for HPFH or $\delta\beta$ -thalassaemia. In addition, we unambiguously demonstrate direct physical linkage between the fetal and adult globin genes.

Deletion mapping

Deletions in two cases each of homozygous hereditary persistence of fetal haemoglobin (HPFH-1 and 2)^{3,5,11,12} and of homozygous $\delta\beta$ -thalassaemia ($\delta\beta$ -1 and 2)^{9,10,22} were mapped relative to restriction endonuclease cleavage sites within and surrounding the γ -, δ - and β -globin genes in normal human DNA. The map of these cleavage sites in normal DNA shown in Fig. 1a was derived by analysis of cloned globin genes¹⁹ and by blot hybridisation experiments described here and elsewhere^{18,20,24}. To scan the entire β -like gene locus systematically for the location of deletions, blots of each type of mutant DNA were first hybridised individually with *in vitro* labelled β - or γ -globin cDNA plasmids²⁵ or with a cloned DNA fragment from a region between the adult and fetal globin genes. Precise localisation of the end points of the deletions was then accomplished by additional enzyme digestions as described in detail below.

HPFH-1 DNA

HPFH-1 DNA was obtained from an individual homozygous for a type of HPFH in which both the $\Lambda\gamma$ and $\Gamma\gamma$ polypeptides are produced in adults³. Initially, normal and HPFH-1 DNA were compared by blot hybridisation procedures using the β -globin cDNA plasmid as probe (Fig. 2a). As expected from the restriction map of Fig. 1a, digestion of normal DNA with *Eco*RI produces four fragments (5.2, 3.6, 2.3 and 1.75 kilobases) which hybridise the β -globin cDNA probe (Fig. 2a). Similarly, digestion with *Taq*I produces two hybridising fragments (8.5 and 3.0 kilobases) while digestion with *Xba*I produces a single hybridising fragment (10.8 kilobases). In contrast, no hybridising fragments are observed following digestion of HPFH-1 DNA, indicating that, in agreement with liquid hybridisation³ and blot hybridisation¹⁰ experiments, HPFH-1 DNA has little or no β - or δ -globin gene sequences.

Fragments characteristic of normal DNA were observed when HPFH-1 DNA was digested with *Eco*RI, *Bgl*II or *Sac*I and hybridised to the γ -globin probe (Fig. 2b). These results indicate that the left end point of the deletion in HPFH-1 DNA is located to the left of the δ -globin gene as shown in Fig. 1a. We therefore hybridised blots of HPFH-1 DNA with a cloned 0.5-kilobase fragment which lies 4 kilobases to the left of the δ -globin gene¹⁹ ('RI H'; see Fig. 1a) to determine whether the RI H sequence is present in HPFH-1 DNA and to map the surrounding restriction enzyme sites. As shown in Fig. 3a, a 7.2-kilobase fragment is observed in both normal and HPFH-1 DNA digested with *Eco*RI, indicating that the RI H sequence is present and that the end point of the HPFH-1 deletion is located to the right of RI H.

To map this end point further, additional enzyme digests were carried out. Digestion of normal DNA with *Pst*I, *Xba*I or *Taq*I yields single fragments of 4.8, 1.2 and 2.0 kilobases, respectively, which hybridise to the RI H probe (Fig. 3b). Double digests with *Eco*RI were carried out to map these sites in normal DNA with respect to RI H. Fragments of 3.2 kilobases (*Pst*I/*Eco*RI), 0.5 kilobases (*Xba*I/*Eco*RI), and 1.6 kilobases (*Taq*I/*Eco*RI) were observed. These results identify *Pst*I, *Xba*I and *Taq*I cleavage sites 3.2, 0.5 and 1.6 kilobases, respectively, to the left of the *Eco*RI site at the right end of RI H. By subtraction from the size of the single digest products, *Pst*I, *Xba*I and *Taq*I sites lie 1.6, 0.7 and 0.4 kilobases to the right of this *Eco*RI site (see Fig. 1a).

Similar single and double enzyme digests were carried out with HPFH-1 DNA to determine which of these sites is altered (Fig. 3b). The sizes of the RI H hybridising fragments in the double enzyme digests are identical to those observed with normal DNA, indicating that the *Pst*I, *Xba*I and *Taq*I sites to the left of RI H are not altered in HPFH-1 DNA. However, the sizes of all the hybridising fragments produced in single enzyme digests are altered, indicating that sites to the right of RI H are removed by the deletion (Fig. 3b). These data place the end

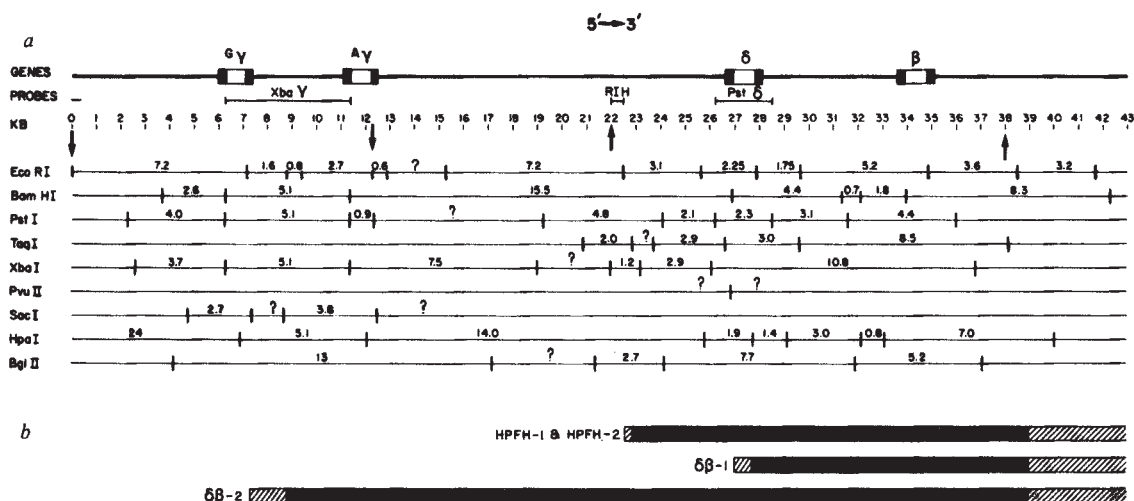


Fig. 1 A map of restriction endonuclease cleavage sites within and surrounding the human γ -, δ - and β -globin genes. *a*, The arrangement of fetal and adult β -like globin genes within a 43-kilobase segment of human DNA. Physical linkage of the adult δ - and β -globin genes^{18,19} and the fetal $\Gamma\gamma$ - and $\Lambda\gamma$ -globin genes²⁰ has been previously described. The physical linkage of the fetal and adult globin genes is described in the text. The direction of transcription of the four linked genes is left to right (5' $\rightarrow$ 3'). The solid boxes represent the locations of mRNA coding regions. The open boxes represent the large non-coding intervening sequence in each gene^{18-20,23,24}. This intervening sequence is located between codons 104 and 105 in the δ , β (ref. 19) and $\Lambda\gamma$ (ref. 23) genes. Sizes of the $\Gamma\gamma$ -, $\Lambda\gamma$ -, δ - and β -globin intervening sequences are approximately 850, 850, 950 and 900 base pairs, respectively. A smaller intervening sequence identified in the β - and $\Lambda\gamma$ -genes and presumably present in the $\Gamma\gamma$ - and δ -genes is not shown. The brackets below the gene map indicate the positions of the *Xba*I, *RI H* and *Pst*I δ fragments. These fragments are used as hybridisation probes in some experiments and are described in detail in the appropriate figure legends. For each of the restriction endonucleases indicated on the left, the relative locations of cleavage sites are marked by a vertical line. The sizes for the restriction enzyme fragments are given in kilobases. Question marks indicate that the presence of additional cleavage sites has not been determined. The locations of cleavage sites within the region delineated by the upward-pointing arrows were determined from an analysis of the clone H β G1 which contains the linked δ - and β -globin genes (ref. 19 and our unpublished results). Some cleavage sites within the region delineated by the downward-pointing arrows were determined by an analysis of the clone H γ G1 which contains the $\Gamma\gamma$ - and $\Lambda\gamma$ -globin genes (unpublished results). Additional cleavage sites within and surrounding the γ -, δ - and β -globin genes were mapped by blot hybridisation experiments (refs 18, 20, and this paper). *b*, The regions of the β -like globin gene locus deleted in HPFH and $\delta\beta$ DNAs are indicated by solid boxes. The precise locations of the end points of the deletions are within the regions specified by the hatched boxes. The locations of the rightward ends of the deletions are not known.

point of the HPFH-1 deletion within the 400-base pair region between the *EcoRI* site bounding the RIH fragment and the *TaqI* site to the right. The deletion therefore extends from approximately 4 kilobases to the left of δ to at least 4 kilobases beyond the 3' end of the β -globin gene, which is the rightward limit of our restriction map. Thus, the deletion in HPFH-1 DNA removes at least 16 kilobases of DNA (Fig. 1b).

HPFH-2 DNA

DNA from a second case of homozygous HPFH was also examined¹⁰⁻¹². This DNA was obtained from an individual apparently unrelated to the HPFH-1 case. As with HPFH-1, HPFH-2 lacks the δ - and β -globin genes and contains no detectable alteration within or closely surrounding the γ -globin genes (data not shown). When HPFH-2 DNA was examined in an experiment identical to that of Fig. 3b, an identical result was obtained (data not shown). Therefore, the deletion in HPFH-2 DNA is indistinguishable from that in HPFH-1 DNA.

$\delta\beta$ -1 DNA

Hybridisation of $\delta\beta$ -1 DNA with the γ -globin probe revealed fragments characteristic of normal DNA (Fig. 2b). However, with the β -globin probe (Fig. 2a) a single hybridising fragment was observed in $\delta\beta$ -1 DNA digested with *EcoRI* (3.0 kilobases) (see also ref. 9), *TaqI* (2.3 kilobases) and *XbaI* (2.5 kilobases). All these sizes differ from those found in normal DNA. This result indicates that a portion of either the β - or δ -globin gene is

present in $\delta\beta$ -1 DNA. Hybridisation of *BamHI*-digested $\delta\beta$ -1 DNA with the RIH probe revealed a fragment identical to that found in normal DNA (data not shown). This result suggests that $\delta\beta$ -1 DNA might contain 5'- δ -gene sequences, as the *BamHI* site nearest to the RIH fragment is located within the δ -gene. The restriction enzyme *PvuII* was used to demonstrate the presence of 5'- δ -sequences of $\delta\beta$ -1 DNA. A *PvuII* site is present in the 5'-end of the δ -globin gene and no other *PvuII* sites are found in either the β - or δ -globin genes¹⁸. To detect δ -globin gene sequences we used a cloned *PstI* fragment ('Pst δ '; see Fig. 1a) which contains the δ -globin gene and surrounding sequences.

Digestion of $\delta\beta$ -1 DNA with *PstI* produces a single fragment of 10.5 kilobases which hybridises to the *PstI* δ probe. Fragments of 7.3 and 0.9 kilobases are produced in a *PstI*/*PvuII* double digest (Fig. 4a). The presence of a fragment which is cut by *PvuII* to produce two fragments which hybridise the δ -globin probe indicates that a portion of the δ -globin gene is retained in $\delta\beta$ -1 DNA. A similar conclusion was obtained from *EcoRI* and *PvuII* digests (data not shown).

A *BamHI* site is located to the right of the *PvuII* site in the δ -globin gene. To determine whether this *BamHI* site is present in $\delta\beta$ -1 DNA, double digests with *BamHI* were carried out. A *BamHI*/*PstI* digest of normal DNA produces fragments of 0.95 and 1.35 kilobases which hybridise the *PstI* δ probe (Fig. 4b; a weakly hybridising β -gene fragment is also seen). The 0.95-kilobase fragment corresponds to the *PstI*/*BamHI* fragment to the left of the intragenic *BamHI* site, whereas the 1.35-kilobase fragment corresponds to the *BamHI*/*PstI* fragment to the right of this site. A *BamHI*/*PstI* digest of $\delta\beta$ -1 DNA produces a 0.95 and a 3.3-kilobase fragment but no 1.35-kilobase fragment. Similarly, a *BamHI*/*EcoRI* digest of normal DNA produces fragments of 1.3, 1.0 and 1.75 kilobases (plus a β -gene fragment). The 1.3-kilobase fragment corresponds to the *EcoRI*/*BamHI* fragment to the left of the intragenic *BamHI* site; the 1.0-kilobase fragment to the *BamHI*/*EcoRI* fragment containing the δ -globin intervening sequence; and the 1.75-kilobase fragment to the *EcoRI* fragment containing the 3'-portion of the δ -globin gene. A *BamHI*/*EcoRI* digest of $\delta\beta$ -1 DNA produces a 1.3-kilobase 5'- δ -gene fragment and a 1.9-kilobase fragment but not the 1.0- or 1.75-kilobase fragments of normal DNA. Together these results indicate that the restriction enzyme sites leftward from the intragenic *BamHI* site are present and that the *EcoRI* and *PstI* sites to the right of the *BamHI* site are missing in $\delta\beta$ -1 DNA. This conclusion was confirmed by comparing *EcoRI*/*PvuII* or *PstI*/*PvuII* digests of normal and $\delta\beta$ -1 DNAs (data not shown).

To further delineate the end point of the deletion, we compared the *HpaI* digestion patterns of normal and $\delta\beta$ -1 DNA, as an *HpaI* site is located within the intervening sequence of the δ -globin gene approximately 150 base pairs to the left of the *EcoRI* site (unpublished results). Digestion of normal DNA with *HpaI* produces fragments of 1.9 and 1.4 kilobases which hybridise with the δ -globin probe (Fig. 4c; a large β -gene fragment is also observed). The 1.9-kilobase fragment contains the δ -globin gene and the portion of the intervening sequence to the left of the *HpaI* site, and the 1.4-kilobase fragment contains sequences to the right of the *HpaI* site. Digestion of $\delta\beta$ -1 DNA yields only one hybridising fragment of 2.1 kilobases, indicating that at least two of the three *HpaI* sites which delineate the 1.9- and 1.4-kilobase fragments in normal DNA are missing. Additional experiments using the RIH probe indicate that the *HpaI* site to the left of the δ -globin gene in normal DNA is present in $\delta\beta$ -1 DNA (data not shown). Therefore, the *HpaI* site near the right end of the intervening sequence is absent in $\delta\beta$ -1 DNA.

In double digests with *BamHI* and either *PstI* or *EcoRI*, sequences to the right of the *BamHI* site hybridised to the *PstI* δ probe. To form a stable hybrid in the conditions used, more than the 12 base pairs of coding sequence between the *BamHI* site and the beginning of the intervening sequence are required. Hence, at least some of the δ -globin intervening sequence must be present in $\delta\beta$ -1 DNA. Therefore, the deletion in $\delta\beta$ -1 DNA

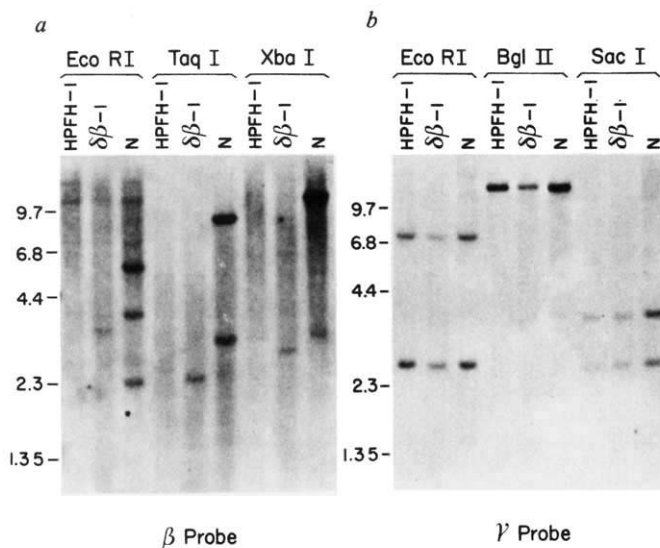


Fig. 2 Hybridisation of normal, HPFH-1 and $\delta\beta$ -1 DNAs with cloned β - and γ -globin cDNA probes. Human DNA was isolated from transformed lymphoid cells^{3,9} using the procedure of Blin and Stafford²⁶. DNA samples were digested separately with *EcoRI*, *TaqI* or *XbaI*, electrophoresed in a horizontal 0.8% agarose slab gel and transferred to nitrocellulose filter paper¹³. The *HhaI* fragments containing β - or γ -globin cDNA sequences were isolated from the plasmids pJW102 or pJW151 (ref. 25), respectively, labelled with ³²P by nick translation²⁷ and hybridised to the filter-bound DNA^{13,15}. (The recombinant DNA containment level of P3⁺EK2 was used as outlined in the National Institutes of Health Recombinant DNA Research Guidelines.) Following hybridisation the filters were washed sequentially in 1×, 1/2× and 1/4× washing buffer at 68°C (1× buffer is 50 mM NaCl, 10 mM Tris-HCl, pH 7.5, 1 mM EDTA, 0.1% SDS) and exposed to X-ray film using two DuPont Cronex lighting plus intensifier screens. Bacteriophage λ DNA digested with *HindIII* (ref. 19) and pBR322 DNA digested with *HinfI* (ref. 28) were included in the gel as size standards. The positions of some of the markers are indicated to the left of the autoradiogram (in kilobase pairs). *a*, Digests of $\delta\beta$ -1, HPFH-1 and normal (N) DNA hybridised with the β -globin probe. The position of the weakly hybridising 1.75-kilobase 3'- δ *EcoRI* fragment is indicated by the dot. *b*, Digests of the same DNAs hybridised with the γ -globin probe. Two additional *EcoRI* fragments (1.6 and 0.6 kilobases)²⁰ are not detected by the γ -globin cDNA plasmid used here, which lacks a portion of the 3'-mRNA sequence.

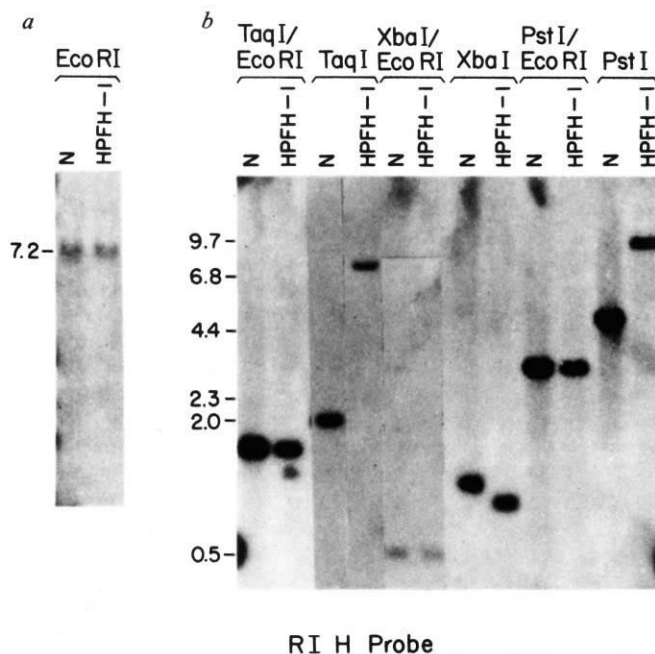


Fig. 3 Hybridisation of normal and HPFH-1 DNA with the RI H fragment. Normal and HPFH-1 DNAs were hybridised as in Fig. 2 with *in vitro* labelled RI H DNA. The RI H fragment in H β G1 DNA is located approximately 4 kilobases to the left of the δ -globin gene (Fig. 1a) and is bounded on the right end by an *Eco*RI site present in genomic DNA and on the left end by an artificial *Eco*RI site created during the construction of the human DNA library¹⁹. The RI H fragment was isolated from H β G1, subcloned into the *Eco*RI site of the plasmid pMB9 (ref. 29) and grown in *Escherichia coli* χ 1776 (ref. 30). The RI H fragment was isolated from purified plasmid DNA of the subclone following *Eco*RI digestion. *a*, *Eco*RI digest; *b*, single and double enzyme digests using *Taq*I, *Xba*I and *Pst*I with *Eco*RI.

begins within the large intervening sequence of the δ -globin gene and extends to at least 4 kilobases to the right of the β -globin gene (Fig. 1b).

$\delta\beta$ -2 DNA

When $\delta\beta$ -2 DNA¹⁰ was digested with various restriction endonucleases, blotted and hybridised with the β -globin probe, no hybridisation was observed (ref. 10 and our unpublished results). Moreover, we detect no hybridisation with the RI H probe. The deletion in $\delta\beta$ -2 DNA therefore extends through the δ - and β -globin genes and beyond the RI H sequence. We next determined whether the deletion also covers restriction enzyme sites near the γ -globin genes. In *Bgl*II digests of normal DNA a 13-kilobase fragment containing both γ -globin genes was detected (Fig. 5a). However, in $\delta\beta$ -2 DNA a 7-kilobase fragment was observed, indicating that one of the two *Bgl*II sites surrounding the γ -globin genes is deleted in $\delta\beta$ -2 DNA. Digestion of normal DNA with *Eco*RI produces fragments of 7.2 and 2.7 kilobases which contain the 5'-portions of the $\alpha\gamma$ and $\beta\gamma$ -genes, respectively²⁰. (Two fragments of 1.7 and 0.8 kilobases containing the 3' portions of these genes²⁰ are not efficiently detected with the pJW151 γ cDNA probe.) Only the 7.2-kilobase 5' $\alpha\gamma$ fragment was detected in $\delta\beta$ -2 DNA. Thus, part or all of the $\alpha\gamma$ -gene is deleted in $\delta\beta$ -2 DNA and sequences to the left of and including the intragenic *Eco*RI site in the $\alpha\gamma$ -gene are present. This conclusion is consistent with the *Bam*HI hybridisation pattern. *Bam*HI digestion of normal DNA produces fragments of 15.5, 5.1 and 2.6 kilobases which hybridise to the γ -globin probe (Fig. 5b). These fragments contain, respectively, 3' $\alpha\gamma$, 3' $\alpha\gamma$ plus 5' $\alpha\gamma$, and 5' $\alpha\gamma$ sequences. Only the 5' $\alpha\gamma$ fragment is observed in *Bam*HI digests of $\delta\beta$ -2 DNA. The end point of the deletion within the

γ -globin gene locus was more precisely mapped using the enzyme *Sac*I.

*Sac*I digestion of normal DNA produces fragments of 3.8 and 2.7 kilobases which hybridise the γ -globin probe (Fig. 5b). These fragments contain the $\alpha\gamma$ - and $\beta\gamma$ -gene sequences, respectively. *Sac*I digestion of $\delta\beta$ -2 DNA produces only the 2.7-kilobase $\beta\gamma$ fragment, indicating that the *Sac*I site to the right of the $\alpha\gamma$ -intragenic *Eco*RI site is present in $\delta\beta$ -2 DNA. No additional fragments were observed, again implying that part or all of the $\alpha\gamma$ -gene is deleted.

To determine whether the *Eco*RI site to the right of the $\alpha\gamma$ -gene is present in $\delta\beta$ -2 DNA, a cloned genomic DNA fragment containing 3'- $\alpha\gamma$ sequences was used as a probe (*Xba* γ ; Fig. 1). Figure 5c shows the *Eco*RI lanes from the filter of Fig. 5a reprobbed with the *Xba* γ fragment. With the *Xba* γ probe the 1.7-kilobase 3'- $\alpha\gamma$ *Eco*RI fragment of normal DNA is detected (compare Fig. 5a and c). This fragment is missing in $\delta\beta$ -2 DNA, indicating that the deletion removes the *Eco*RI site to the right of the $\alpha\gamma$ -gene. Thus, the deletion in $\delta\beta$ -2 DNA begins in the region between the *Sac*I site within the $\alpha\gamma$ -gene³² and the *Eco*RI site to the right of the gene and extends through the $\alpha\gamma$ -gene. The same (or possibly a second) deletion extends through the RI H, δ - and β -globin sequence (Fig. 1b).

Physical linkage of γ -, δ - and β -globin genes

Linkage between the fetal and adult globin genes was suggested by structural analysis of Hb Kenya¹⁷, a fusion protein consisting of $\alpha\gamma$ -globin N-terminal and β -globin C-terminal sequences. However, direct physical linkage has not been demonstrated. We have used cloned genomic DNA hybridisation probes and the deletion map of HPFH-1 DNA (Fig. 1b) to demonstrate unambiguously physical linkage between the $\alpha\gamma$ - and δ -globin genes.

Initially, we identified large restriction enzyme fragments in normal DNA which hybridise the γ -globin and RI H probes. As shown in Fig. 6a the RI H probe hybridises with a 15.5-kilobase *Bam*HI fragment and a 14.0-kilobase *Hpa*I fragment in normal

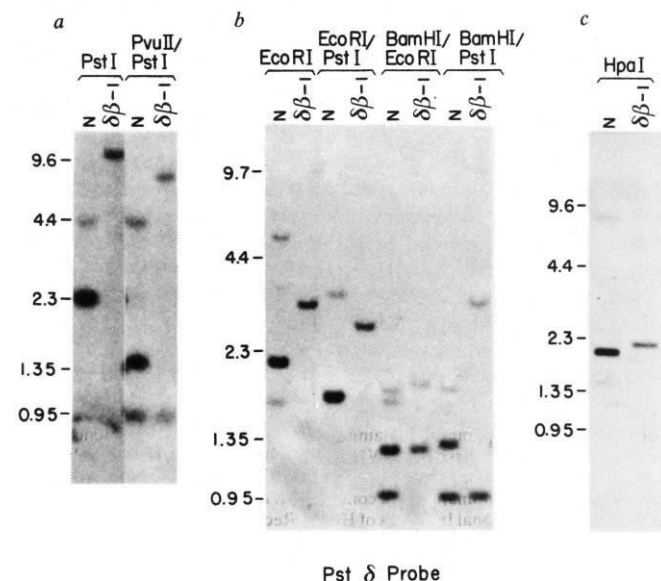


Fig. 4 Hybridisation of normal and $\delta\beta$ -1 DNAs with the δ -globin *Pst*I fragment. Normal and $\delta\beta$ -1 DNAs were hybridised as described in the legend to Fig. 2 with *in vitro* labelled *Pst* δ DNA. The *Pst* δ fragment of H β G1 DNA contains the δ -globin gene and some surrounding sequences (Fig. 1a). The *Pst* δ fragment was isolated from H β G1, subcloned into the *Pst*I site of the plasmid pBR322 (ref. 31) and grown in *E. coli* χ 1776 (ref. 30). The *Pst* δ fragment was isolated from the purified plasmid DNA of the subclone following *Pst*I digestion. *a*, *Pst*I/*Pvu*II double digest; *b*, double digests using *Eco*RI, *Pst*I and *Bam*HI; *c*, *Hpa*I digest.

DNA. Fragments of these sizes are also detected using the pJW151 γ -gene probe (Fig. 6b) or a probe specific for the 3'-end of the γ -genes (Fig. 6c). The 15.5-kilobase fragment corresponds to a 15-kilobase fragment which was previously shown to contain 3'- γ -sequences²⁰. Our mapping experiments (unpublished) indicate that the 14.0-kilobase *HpaI* fragment also contains 3'- γ -sequences. These results indicate that RI H and 3'- γ -sequences are present on fragments of identical size in *BamHI* or *HpaI* digests. Thus, the 3' end of the γ -gene seems to be linked to the RI H fragment. Definitive evidence for this linkage relationship was obtained by comparing the sizes of the *BamHI* and *HpaI* fragments in normal and HPFH-1 DNA using RI H and γ -globin probes. If the linkage arrangement proposed above is correct, the deletion in HPFH-1 should alter the size of the *BamHI* and *HpaI* fragments detected by both the RI H and γ -probes. As predicted, the 15.5-kilobase *BamHI* and 14.0-kilobase *HpaI* fragments of normal DNA are not seen in HPFH-1 DNA, whereas 13.0-kilobase *BamHI* and 13.0-kilobase *HpaI* fragments are observed (Fig. 6a-c). Thus, the deletion in HPFH-1 DNA, which terminates near the RI H fragment, has an identical effect on the size of restriction enzyme fragments detected by either the RI H or γ -gene probes. These results definitely establish physical linkage between the 3'- γ -gene and the RI H fragment.

We have previously shown that the RI H fragment is linked to the 5' end of the δ -globin gene in cloned genomic DNA¹⁹. The linkage of the 3'- γ -gene with the RI H fragment described here, and the linkage of the RI H fragment with the 5'- δ -gene previously described provide unambiguous proof for the linkage arrangement shown in Fig. 1a. Consistent with this conclusion, we have observed that a 15.5-kilobase *BamHI* fragment also hybridises with the β -globin probe (Fig. 6d) in conditions in which cross-hybridisation of γ -gene sequences with the β -globin probe is not observed (compare Fig. 6b and d). The β -like globin genes are therefore arranged in the order: 5'- γ - γ - δ - β 3'.

Discussion

We have used the detailed map of restriction endonuclease cleavage sites within and surrounding the δ - and β -globin genes to locate the end points of deletions in the DNA from individuals with $\delta\beta$ -thalassaemia and HPFH we have designated $\delta\beta$ -1 and HPFH-1 and 2. The deletions in HPFH-1 and 2 are indistinguishable. In addition, the haematological findings in the two cases are virtually identical^{3,4,11,12}. Although there is no known family relationship between the HPFH-1 and 2 individuals, it is possible that the HPFH chromosomes in the two cases had the same ultimate origin. Alternatively, it is possible that the end point of the deletions represents a mutational hot spot so that the two deletions were in fact the products of independent events.

The haematological findings in the $\delta\beta$ -1 and HPFH-1 and 2 cases reveal several significant differences^{3,4,11,12,22}. In $\delta\beta$ -1, characteristics of thalassaemia which include anaemia and hypochromic, microcytic red cells were observed. In addition, analysis of heterozygous relatives indicated that 10% of their haemoglobin was HbF containing both γ - and δ -polypeptide chains and the cellular distribution was heterogeneous (30–40% of the cells stain for HbF). In HPFH-1 and 2 none of the characteristics of thalassaemia was observed. In heterozygous relatives approximately 20–30% of the haemoglobin was HbF (both γ and δ) and the cellular distribution was homogenous (100% of the cells stain for HbF). The high levels of HbF found in $\delta\beta$ -thalassaemia heterozygotes may be due in part to the increased proliferation or selective survival of a small population (1%) of 'F-cells' found in the blood of normal individuals⁴. This is not a general phenomenon, however, as in heterozygous β^0 -thalassaemia, which in most cases is a non-deletion type of β -globin deficiency⁹, very little HbF is found in adults. In fact, homozygous β^0 -thalassaemia is a much more severe type of anaemia than homozygous $\delta\beta$ -thalassaemia.

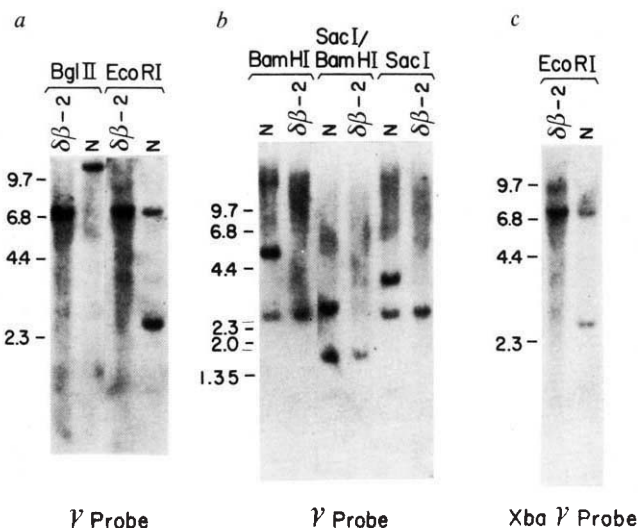


Fig. 5 Hybridisation of normal and $\delta\beta$ -2 DNAs with the cloned γ -globin cDNA plasmid and with the Xba γ fragment. Normal and $\delta\beta$ -2 DNAs were hybridised as in Fig. 2 with the *in vitro* labelled *HhaI* fragment of pJW151 DNA²⁵ or the Xba γ fragment. The Xba γ fragment was isolated following *XbaI* digestion of HyG1, a clone of human DNA containing the γ - and δ -globin genes (unpublished results). The Xba γ fragment contains the DNA sequences between *XbaI* sites located at similar positions in the intervening sequence of the γ - and δ -genes (Fig. 1a). a, *BglII*, *EcoRI* digests using the γ probe; b, *BamHI*, *SacI* single and double digests using the γ probe; c, *EcoRI* lanes from a reprobed with the Xba γ fragment.

Thus, there seems to be a correlation between the extent of the deletion of the δ - and β -globin genes and the level of γ -globin gene expression in adults, suggesting that sequences involved in the suppression of γ -gene expression are located in the δ - β -globin gene region. This deletion model for γ -gene expression in adults was first elaborated by Huisman *et al.*⁷. A comparison of the deletions in HPFH-1 and 2 and $\delta\beta$ -1 suggests that one such regulatory sequence may be located within a 4-kilobase region on the 5' side of the δ -globin gene. An alternative possibility is that HPFH-1 and 2 are double mutants, the second lesion affecting a negative regulatory sequence located in the γ -globin gene region³. Our data rule out the presence of a large deletion of sequences surrounding the γ -genes, but a small deletion or point mutation would not have been detected. Although the occurrence of a second mutation in all the different types of HPFH described seems unlikely, the double mutation hypothesis remains a formal possibility. Furthermore, the possibility that the HPFH phenotype results from the deletion of sequences 3' to the β -globin gene cannot be ruled out by our data since we have not located the rightward end points of the various deletions. However, the fact that individuals with Hb Kenya express the HPFH phenotype argues against this possibility. In Hb Kenya the region between γ and δ genes is thought to be missing but the sequences which lie to the 3' side of the β gene are presumed to be intact⁷.

If there are regulatory sequences surrounding the δ - and β -globin genes, what is the nature of their action? Analysis of heterozygous cases of HPFH in which only one of the γ -globin chains is expressed in adults have led to the suggestion that the HPFH deletion affects γ -gene expression in *cis*. For example, in heterozygotes of γ -HPFH, only the γ -polypeptide is detected in adults even though the δ -chain on the normal chromosome is functional (the δ -polypeptide is detected at birth in individuals with γ -HPFH but disappears during postnatal development)³³. If the regulatory sequence acted in *trans*, sequences on the normal chromosome would suppress the expression of both γ -genes in the heterozygote. An analogous argument can be made for the heterozygous γ - δ -types of HPFH. If, indeed, a

cis-acting regulatory sequence is present in $\delta\beta$ -1 but absent in HPFH-1 and 2 DNAs, its location, 12 kilobases to the 3' side of the γ -gene (Fig. 1), is unprecedented. We know of no example of a 3'-*cis* acting control sequence in either a prokaryotic or eukaryotic gene system. However, it is reasonable to suppose that sequences which lie to the 3' side of a gene could function in globin RNA processing or splicing³⁴, in the activation of a chromosome region similar to the phenomenon of puffing in *Drosophila* polytene chromosomes³⁵, or in an immunoglobulin-type rearrangement of DNA sequences during development³⁶. The availability of cloned segments of DNA covering the entire β -like globin gene locus will make it feasible to test these possibilities.

The correlation of the haematological data of the $\delta\beta$ -2 case with the deletion map of Fig. 1 seems to argue against the existence of a negative regulatory sequence on the 5' end of the δ -globin gene. Clinical data reveal a moderately severe anaemia which includes marked disorders in red cell morphology. The heterozygous parents of $\delta\beta$ -2, who are first cousins, have 10–13% HbF which is non-uniformly distributed. Although peptide analysis of the HbF in $\delta\beta$ -2 has not been reported, our data show that only the γ -chain could be produced. Thus, $\delta\beta$ -2 seems to be an example of $\gamma\delta$ -thalassaemia. However, the deletion model would predict an HPFH phenotype as the putative regulatory sequence located near the 5' end of the δ -globin gene is missing in $\delta\beta$ -2 DNA. This discrepancy might be explained by the fact that the deletion in $\delta\beta$ -2 DNA extends through the γ -gene and into the region between the γ and γ -genes. Only one case each of homozygous γ -HPFH^{37,38} and $\gamma\delta$ -thalassaemia¹⁰ ($\delta\beta$ -2) has been reported. These individuals were designated HPFH or $\delta\beta$ -thalassaemia on the basis of analysis of heterozygous relatives even though both homozygotes displayed many of the characteristics of $\delta\beta$ -thalassaemia. Altay *et al.* have previously discussed⁸ the difficulty in discriminating between γ -types of HPFH and $\delta\beta$ -thalassaemia. Even if γ -gene suppression was eliminated in γ -HPFH, a chain

imbalance resulting from the deletion of the γ -gene and surrounding sequences would result in thalassaemia traits. Thus, it may not be valid to compare γ - and $\gamma\delta$ -types of HPFH. Until further information regarding the locations of deletions in cases of γ -HPFH and $\delta\beta$ -thalassaemia can be obtained, the significance of the $\delta\beta$ -2 result with regard to the deletion model cannot be assessed.

The experiments described here represent an initial step towards localising sequences which are involved in controlling the switch from fetal to adult β -like globin gene expression in man. The construction of a detailed linkage map of the β -like globin genes and the use of cloned hybridisation probes which span the entire region have established the feasibility of mapping deletions in human DNA. Hopefully, a similar analysis of additional cases of HPFH and $\delta\beta$ -thalassaemia will ultimately lead to the identification of sequences which are involved in controlling these developmentally regulated genes.

We thank Dr Bernard Forget for human DNA samples and helpful discussions, Catherine O'Connell for preparing the *Pst* δ and RI H subclones, Stuart Orkin for providing $\delta\beta$ -2 DNA, H. Lazarus for providing transformed lymphocytes from the HPFH-1 case, S. Charache for providing peripheral blood from the HPFH-2 case, Walter Schroeder for useful discussions, and P. Little, R. Flavell and R. Williamson for communicating their results before publication. This work was supported by a grant from the NIH to T.M. and by funds from a Biomedical Research Support Grant to the California Institute of Technology. R.M.L. and E.F.F. were supported by postdoctoral fellowships from the American Cancer Society and the Damon Runyon-Walter Winchell Cancer Fund, respectively. T.M. is a recipient of a Rita Allen Foundation Career Development Award.

Received 5 March; accepted 30 April 1979.

- Weatherall, D. J. & Clegg, J. B. in *The Thalassemia Syndrome* 2nd edn (Blackwell, Oxford, 1972).
- Bunn, F. H., Forget, B. G. & Ranney, H. M. in *Human Hemoglobins* (Saunders, Philadelphia, 1977).
- Forget, B. G. *et al.* *Cell* **7**, 323–329 (1976).
- Otolenghi, S. *et al.* *Cell* **9**, 71–80 (1976).
- Kan, Y. W., Holland, J. P., Dozy, A. M., Charache, S. & Kazanian, H. H. *Nature* **258**, 162–163 (1975).
- Ramirez, F. *et al.* *Nature* **263**, 471–475 (1976).
- Huisman, T. H. J. *et al.* *Ann. N.Y. Acad. Sci.* **232**, 107–124 (1974).
- Altay, C., Schroeder, W. A. & Huisman, T. H. J. *Am. J. Hemat.* **3**, 1–14 (1977).
- Mears, J. G. *et al.* *Proc. natn. Acad. Sci. U.S.A.* **75**, 1222–1226 (1978).
- Orkin, S. H. *et al.* *New Engl. J. Med.* **299**, 166–172 (1978).
- Conley, C. L., Weatherall, D. J., Richardson, S. N., Shepard, M. K. & Charache, S. *Blood* **21**, 261–281 (1963).
- Charache, S. & Conley, C. L. *Ann. N.Y. Acad. Sci.* **165**, 37–41 (1969).
- Southern, E. M. *J. molec. Biol.* **93**, 503–517 (1975).
- Botchan, M., Topp, W. & Sambrook, J. *Cell* **9**, 269–287 (1976).
- Jeffreys, A. J. & Flavell, R. A. *Cell* **12**, 429–439 (1977).
- Baglioni, C. *Proc. natn. Acad. Sci. U.S.A.* **48**, 1880–1886 (1962).
- Huisman, T. H. J., Wrightstone, R. N., Wilson, J. B., Schroeder, W. A. & Kendall, A. G. *Archs Biochem. Biophys.* **153**, 850–853 (1972).
- Flavell, R. A., Kooter, J. M., DeBoer, E., Little, P. F. R. & Williamson, R. *Cell* **15**, 25–41 (1978).
- Lawn, R. M., Fritsch, E. F., Parker, R. C., Blake, G. & Maniatis, T. *Cell* **15**, 1157–1174 (1978).
- Little, P. F. R., Flavell, R. A., Kooter, J. M., Annison, G. & Williamson, R. *Nature* **278**, 227–231 (1979).
- Maniatis, T. *et al.* *Cell* **15**, 687–701 (1978).
- Silvestroni, E., Bianco, I. & Reitano, G. *Acta haemat.* **40**, 220–229 (1968).
- Smithies, O. *et al.* *Science* **202**, 1284–1289 (1978).
- Mears, J. G., Ramirez, F., Leibowitz, D. & Bank, A. *Cell* **15**, 15–23 (1978).
- Wilson, J. T. *et al.* *Nucleic Acids Res.* **5**, 563–581 (1978).
- Blin, N. & Stafford, D. W. *Nucleic Acids Res.* **3**, 2303–2308 (1976).
- Maniatis, T., Jeffrey, A. & Kleid, D. *Proc. natn. Acad. Sci. U.S.A.* **72**, 1184–1188 (1975).
- Sutcliffe, J. G. *Nucleic Acids Res.* **5**, 2721–2728 (1978).
- Rodriguez, R. L., Bolivar, F., Goodman, H. M., Boyer, H. W. & Betlach, M. in *Molecular Mechanisms in the Control of Gene Expression* (eds Nierlich, D. P., W. J. & Fox, C. F.) 471–478 (Academic, New York, 1976).
- Curtis, R. III, Inoue, M., Pereira, D., Hsu, J. C., Alexander, L. & Rock, L. in *Molecular Cloning of Recombinant DNA* (eds Scott, W. A. & Werner, R.) 99–114 (Academic, New York, 1977).
- Bolivar, F. *et al.* *Gene* **2**, 95–113 (1977).
- Forget, B. *et al.* in *Eucaryotic Gene Expression* (eds Axel, R., Maniatis, T. & Fox, C. F.) (Academic, New York, in the press).
- Huisman, T. H. J. *et al.* *Pediat. Res.* **9**, 1–6 (1975).
- Tilghman, S. M., Curtis, P. J., Tiemeier, D. C., Leder, P. & Weissmann, C. *Proc. natn. Acad. Sci. U.S.A.* **75**, 1309–1313 (1978).
- Ashburner, M. *Chromosoma* **21**, 398–428 (1967).
- Tonegawa, S., Brack, C., Hozumi, N. & Pirrotta, V. *Cold Spring Harb. Symp. quant. Biol.* **42**, 921–931 (1977).
- Sukumaran, P. K. *et al.* *Br. J. Haemat.* **23**, 403–417 (1972).
- Ringelmann, B. *et al.* *Biochem. Genet.* **15**, 1083–1096 (1977).

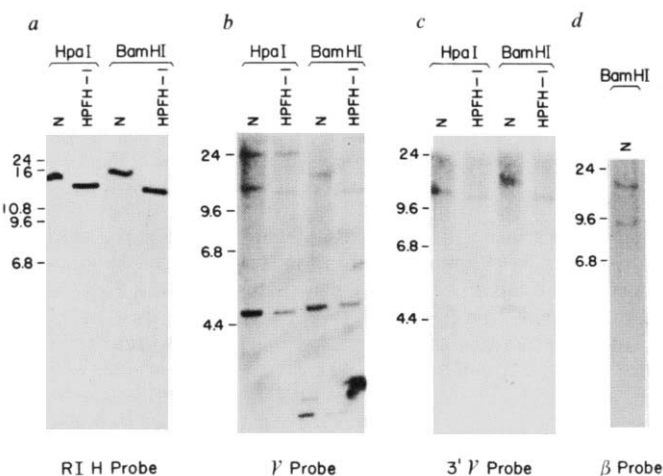


Fig. 6 Hybridisation of normal and HPFH-1 DNA with cloned β - and γ -globin cDNA and the RIH fragment—linkage of fetal and adult β -like globin genes. Normal and HPFH-1 DNAs were hybridised as described in the legend to Fig. 2 with the hybridisation probes indicated below each panel. 5'- and 3'- γ -globin specific DNA was prepared by digesting pJW151 DNA²⁵ with *HhaI* and *BamHI* and purifying the fragments containing γ -globin sequences 5' and 3' to the intragenic *BamHI* site. The 3'- γ -globin specific DNA was labelled *in vitro* by nick translation²⁷. A threefold excess of unlabelled 5'- γ -globin specific DNA was included in the hybridisation as competing DNA. The remaining hybridisation probes have been described (see Figs. 2 and 3 legends). Multimeric forms of Φ X174 DNA were included in the gel as size markers¹⁹. *a*, *HpaI* and *BamHI* digests using the RIH probe; *b*, *HpaI* and *BamHI* digests using the γ probe; *c*, *HpaI* and *BamHI* digests using the 3'- γ probe; *d*, *BamHI* digest using the β probe.

Biosynthesis of the major human red cell sialoglycoprotein, glyophorin A, in a continuous cell line

Mikko Jokinen & Carl G. Gahmberg

Department of Bacteriology and Immunology, University of Helsinki, Haartmaninkatu 3, SF 00290 Helsinki 29, Finland

Leif C. Andersson

Transplantation Laboratory, and Department of Pathology, University of Helsinki

During biosynthesis of glyophorin A in K562 cells a precursor is rapidly transferred through the endoplasmic reticulum membrane with the COOH-terminal remaining in the cytoplasm. This is glycosylated within the cell and appears at the cell surface after about 30 min. The biosynthetic pathway resembles that described for viral membrane glycoproteins.

MOST, if not all, surface proteins of mammalian cells are glycoproteins^{1,2}. Their peptide portions often span the lipid bilayer membrane^{3,4} and their carbohydrate is exposed to the external milieu⁵⁻⁷. The biosynthetic assembly of the peptide and sugar portions and the mechanisms leading to the appropriate membrane disposition of these proteins raise intriguing questions.

Studies on virus-coded proteins have provided us with most of our more detailed knowledge of integral membrane protein biosynthesis. The polypeptides are synthesised on membrane-bound ribosomes and transferred as nascent chains through the endoplasmic reticulum membrane; glycosylation takes place on the luminal side⁸⁻¹⁰. The attachment of the core portions of the oligosaccharides may occur, through lipid-linked intermediates, even before the polypeptide chains have been completely synthesised. As the virus proteins use the host cell machinery for their biosynthesis, similar mechanisms may be anticipated for mammalian surface proteins.

Most of our information on cell membrane structure is derived from studies on the easily available human erythrocyte membrane^{2,13,14}. Its major sialic acid-rich glycoprotein, glyophorin A (refs 15, 16), is the best characterised integral membrane protein. Glyophorin A spans the membrane with its NH₂-terminus on the outside and its COOH-terminus in the cytoplasm^{3,4}, and has a molecular weight (MW) of 31,000. On polyacrylamide gel electrophoresis in the presence of SDS it shows a higher apparent MW because of its 60% carbohydrate content^{3,17}. Determination of its amino acid sequence^{18,19} indicates that it has 15 serine/threonine-linked oligosaccharides and one asparagine-linked complex oligosaccharide on the external surface of the erythrocyte.

We have produced specific anti-glyophorin A antiserum²⁰ by immunising rabbits with a crude preparation of glyophorin followed by absorption with En(a-) red cell membranes, which lack glyophorin^{16,21-23}. The availability of this specific reagent for glyophorin A and our extensive knowledge of the molecular structure of this glycoprotein provide a potentially useful system for elucidating the pathways involved in mammalian plasma membrane glycoprotein biosynthesis.

We have previously observed that the human continuous leukaemia cell line K562 (ref. 24) is erythroid, and synthesises and expresses glyophorin A on its surface^{25,26}. In this report we

outline the biosynthesis of glyophorin A from a precursor form, its insertion into the endoplasmic reticulum membrane, its subsequent glycosylation and its migration to the plasma membrane.

Identification of a precursor for glyophorin A

For studies on the biosynthesis of glyophorin A, K562 cells were labelled with ³⁵S-methionine for 5 min and chased with medium containing non-radioactive methionine. The labelled cells were solubilised in buffer containing Triton X-100 and immune precipitation was carried out with anti-glyophorin A antiserum or control serum and protein A-containing staphylococci²⁰. The precipitates were analysed by polyacrylamide slab gel electrophoresis in the presence of SDS followed by fluorography. After 5 min of labelling a single protein was specifically precipitated. It had an apparent MW of 37,000 and was designated GP_a (Fig. 1a, B). The use of internal labelling meant that a high background radioactivity could not be avoided and therefore the appearance of the specifically precipitated protein could only be followed by using controls with preimmunisation serum. When chased for 10 min (Fig. 1a, D) the final glyophorin A molecule (designated GP_c), with an apparent MW of 39,000, became visible, and its concentration relative to that of GP_a increased during chase for 25–60 min (Fig. 1a, F–J).

Appearance of glyophorin A at the cell surface

As surface-exposed glyophorin A can be cleaved by trypsin^{15,26} the time course of its externalisation was followed by trypsin treatment of labelled intact cells. After labelling for 5 min followed by chase for 10 min, neither the precursor protein, GP_a, nor the completed glyophorin A (GP_c) had reached the cell surface, as indicated by their insensitivity to trypsinisation (Fig. 1b, E, G). After 25 and 45 min chase no GP_c was seen in trypsin-treated cells, but GP_a was still detected (Fig. 1b, I, K). In the absence of trypsin treatment the GP_c band was stronger than the GP_a band at these time points (compare Fig. 1a, F, H). This shows that after 25 min glyophorin A had appeared at the cell surface but that GP_a had not become exposed.

Glyophorin A precursor is incompletely glycosylated

We followed the sequence of glycosylation using lectin columns of known specificity. The glyophorin A precursor, GP_a, already contained glucose/mannose-like residues because it bound to lentil lectin-Sepharose and could be eluted with α -methyl mannoside (Fig. 2B). However, the following results indicate that GP_a does not contain significant amounts of galactose or

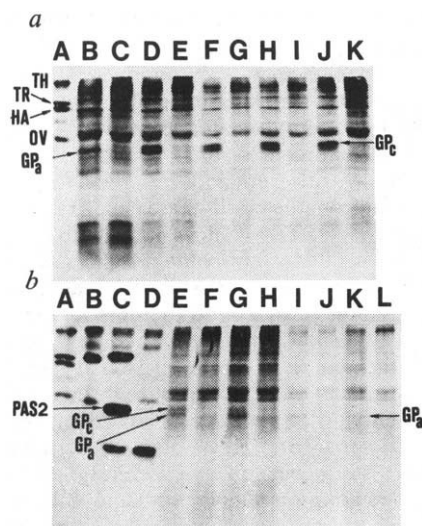


Fig. 1 *a*, Fluorography of a polyacrylamide slab gel of immune precipitates obtained from pulse-chased ^{35}S -methionine-labelled K562 cells with anti-glycophorin A antiserum and control serum. K562 cells (80×10^6) were grown in RPMI-1640 culture medium (GIBCO) containing 10% newborn calf serum. After washing three times with 0.15 M NaCl and 0.01 M sodium phosphate, pH 7.4 (PBS), at 37°C , the cells were suspended in 5 ml of methionine-free Eagle's minimal essential medium (MEM) containing 0.2% bovine serum albumin (Sigma). The cells were incubated with gentle shaking in a CO_2 -atmosphere for 30 min at 37°C and 1.2 mCi of ^{35}S -methionine (Radiochemical Centre, 1,130 Ci mmol $^{-1}$) added. After 5 min 1 ml of the suspension was withdrawn to ice-cold PBS and the cells immediately washed three times with PBS at 0°C . The rest of the pulse-labelled cells were rapidly washed at 37°C with MEM containing a 50-fold excess of cold methionine and suspended in 4 ml of this medium. After incubation for 10, 25, 45 and 60 min 1-ml aliquots were removed and the cells washed at 0°C as above. All washed cell samples were then suspended in 4 ml of PBS at 0°C and 1 ml from each taken for trypsinisation (see *b* below). The rest of the samples were centrifuged and the cells dissolved at 0°C in PBS containing 1% Triton X-100, 1% ethanol and 2 mM phenylmethylsulphonylfluoride (Sigma) as a protease inhibitor (buffer A). Aliquots of the Triton X-100 extracts were taken for immune precipitation experiments using 5 μl anti-glycophorin A antiserum or control serum and the *Staphylococcus aureus* protein A technique as described in detail previously 20,26 . The immune precipitates were studied by polyacrylamide slab gel electrophoresis in the presence of SDS 34 using a 12% acrylamide concentration in the separating gel. The treatment of the slab gels for fluorography 35 and the ^{14}C -labelled 36 standard proteins have been described earlier 37 . A, ^{14}C -labelled standard proteins: TH = thyroglobulin; TR = transferrin; HA = human albumin; OV = ovalbumin; B, immune precipitate obtained with anti-glycophorin A antiserum from K562 cells labelled for 5 min with ^{35}S -methionine; C, with preimmune serum; D, immune precipitate obtained with antiserum from cells labelled for 5 min with ^{35}S -methionine followed by chase for 10 min; E, with preimmune serum; F, immune precipitate obtained with antiserum from cells after 25 min chase; G, with preimmune serum; H, immune precipitate obtained with antiserum from cells after 45 min chase; I, with preimmune serum; J, immune precipitate obtained with antiserum from cells after 60 min chase; K, with preimmune serum. *b*, Fluorography of a polyacrylamide slab gel of labelled erythrocyte membranes and immune precipitates obtained from trypsin-treated K562 cells labelled with ^{35}S -methionine. Part of the ^{35}S -methionine-labelled cells obtained from the experiment described for *a* were treated with 0.1 mg ml $^{-1}$ trypsin (Merck, Darmstadt) for 10 min at 37°C . The samples were then immediately washed with PBS at 0°C and solubilised in buffer A. Immune precipitations and polyacrylamide slab gel electrophoresis were carried out as described in *a*. A, ^{14}C -labelled standard proteins (same as in *a*); B, surface glycoprotein pattern of erythrocyte membranes labelled with ^3H after treatment of erythrocytes with neuraminidase and galactose oxidase followed by NaB^3H_4 (ref. 6); C, surface glycoprotein pattern of erythrocyte membranes obtained after labelling erythrocytes by the periodate/ NaB^3H_4 method 27 ; PAS 2 = predominantly glycophorin A monomer; D, surface glycoprotein pattern of periodate/ NaB^3H_4 -labelled membranes after treatment of labelled intact erythrocytes with 0.1 mg ml $^{-1}$ trypsin for 10 min at 37°C ; E, immune precipitate obtained with anti-glycophorin A antiserum from trypsinised K562 cells that had been labelled with ^{35}S -methionine for 5 min; F, with preimmune serum; G, immune precipitate obtained with antiserum from cells labelled for 5 min with ^{35}S -methionine followed by chase for 10 min; H, with preimmune serum; I, immune precipitate with antiserum from cells after 25 min chase; J, with preimmune serum; K, immune precipitate with antiserum from cells after 45 min chase; L, with preimmune serum.

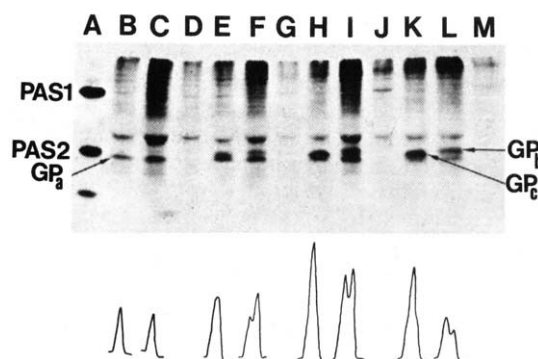


Fig. 2 Fluorography of a polyacrylamide slab gel of glycophorin molecules bound to lentil lectin-Sepharose and the effect of neuraminidase on their electrophoretic mobilities. K562 cells were labelled with ^{35}S -methionine as described in Fig. 1*a* legend, washed at 0°C with PBS, solubilised in buffer A and applied to lentil lectin-Sepharose columns. *Lens culinaris* lectin was prepared by affinity chromatography 38 and coupled to CNBr-activated Sepharose 4B as described in detail previously 37 . The material eluted with 0.1 M α -methyl mannose (Calbiochem) was pooled and concentrated by vacuum dialysis to 3 ml. To 1-ml aliquots from each sample was added 0.3 ml of Dulbecco's PBS containing Ca^{2+} and Mg^{2+} and 100 units of *Vibrio cholerae* neuraminidase (Behringwerke). The neuraminidase preparation did not contain proteolytic activity when assayed as described previously 6 . Two other identical aliquots did not receive neuraminidase. All samples were incubated for 10 min at 37°C . After incubation one control sample and the neuraminidase-treated sample were incubated with 5 μl anti-glycophorin A antiserum and the third tube with preimmune serum followed by protein A-containing staphylococci. The immune precipitates were then analysed by polyacrylamide slab gel electrophoresis as described above. A, surface glycoprotein pattern of periodate/ NaB^3H_4 -labelled erythrocyte membranes; B, immune precipitate obtained with anti-glycophorin A antiserum from the glycoprotein fraction of K562 cells labelled for 5 min with ^{35}S -methionine; C, immune precipitate with antiserum from the neuraminidase-treated glycoprotein fraction of cells labelled for 5 min; D, with preimmune serum; E, immune precipitate obtained with antiserum from the glycoprotein fraction of cells labelled for 5 min followed by 10 min chase; F, immune precipitate with antiserum from the neuraminidase-treated glycoprotein fraction of cells labelled for 5 min followed by 10 min chase; G, with preimmune serum; H, immune precipitate obtained with antiserum after 25 min chase; I, immune precipitate with antiserum of the 25-min chase sample treated with neuraminidase; J, with preimmune serum; K, immune precipitate obtained with antiserum after 60 min chase; L, immune precipitate with antiserum of the 60-min chase sample treated with neuraminidase; M, with preimmune serum. The corresponding scanning patterns of the glycophorin regions are shown below and were obtained with a Joyce-Loebl Chromoscan apparatus.

sialic acids. First, it did not bind to *Ricinus communis* lectin-Sepharose columns specific for β -D-galactose (data not shown). Second, as it is known that neuraminidase treatment of erythrocyte glycophorin A reduces its mobility on SDS-gel electrophoresis this phenomenon was used to determine the presence of sialic acids 27 . Part of the ^{35}S -methionine-labelled eluates from the lentil lectin columns were treated with neuraminidase before immune precipitation. GP_a mobility on SDS-gel electrophoresis did not change after neuraminidase treatment, but instead of final glycophorin A, a molecule, designated GP_b , with an apparent MW of 41,000, was detected. This demonstrates that GP_c contains sialic acids. GP_c could be adsorbed not only to lentil lectin-Sepharose but also to *R. communis* lectin-Sepharose columns, indicating that it contains terminal β -D-galactosyl residues. This experiment also indicates that at this stage GP_a does not contain the penultimate galactosyl residues because its mobility would then have corresponded to that of GP_b .

The NH_2 -terminal portion of glycophorin A precursor is rapidly incorporated into microsomes

The incorporation of the glycophorin A precursor into endoplasmic reticulum was studied by labelling K562 cells for 5 min with ^{35}S -methionine. The labelled cells were then immediately

cooled to 0 °C, homogenised and a microsomal fraction prepared. Half of this fraction was treated with trypsin. The untreated and the trypsin-treated samples were solubilised in buffer A (see Fig. 1a legend) and passed through lentil lectin columns; the glycoproteins were eluted with α -methylmannoside and immune precipitated with anti-glycophorin A antiserum. The immune precipitates from the trypsin-treated and untreated samples were electrophoresed on parallel slots of a polyacrylamide slab gel. Complete, 37,000 MW, glycophorin precursor, GP_a, was recovered from untreated microsomes (Fig. 3B) but the molecule (GP_{at}) obtained from the trypsin-treated microsomes had an apparent MW of 34,000 (Fig. 3C). This shows that a fragment of approximate MW 3,000 could be cleaved off by trypsin at this stage.

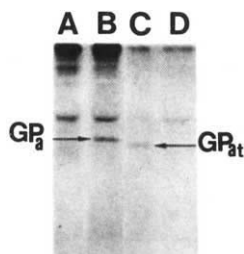


Fig. 3 Fluorography patterns of immune precipitates from K562 microsomes analysed by polyacrylamide slab gel electrophoresis. K562 cells (170×10^6) were labelled for 5 min with ^{35}S -methionine as described in Fig. 1a legend. They were then rapidly washed three times with PBS at 0 °C and suspended for 10 min in 15 ml of 10 mM KCl, 10 mM Tris and 1.5 mM MgCl_2 , pH 7.4, at 0 °C. The cells were then homogenised in a tight-fitting Dounce homogeniser with 20 strokes and the nuclei removed by centrifugation at 1,000g for 10 min. The supernatant was centrifuged at 50,000g for 90 min at 4 °C in a Spinco rotor 30 and the microsomal pellet obtained suspended in 1 ml of Dulbecco's PBS. Half of this material was incubated with 0.01 mg ml⁻¹ trypsin for 10 min at 37 °C. The control sample was handled in the same way except that trypsin was omitted. After incubation 0.5 mg of soybean trypsin inhibitor (Sigma) was added to both samples. These were then dissolved in buffer A, both divided into two and immune precipitated with either 5 μl anti-glycophorin A antiserum or preimmune serum. The immune precipitates were then analysed on a 12% polyacrylamide slab gel. A, Control microsomes + preimmune serum; B, control microsomes + anti-glycophorin A antiserum; C, trypsin-treated microsomes + anti-glycophorin A antiserum; D, trypsin-treated microsomes + preimmune serum. Note the decrease in the apparent MW of trypsin-treated GP_a.

Discussion

We recently showed that glycophorin A is expressed not only on mature red cells but also on nucleated precursor cells in human bone marrow²⁰. This finding prompted us to look for cell lines synthesising glycophorin A. Our observation that the human continuous cell line K562 is erythroid²⁵ was the basis for the present experiments. The cells expressed 10^6 copies per cell of a surface glycoprotein which is indistinguishable from erythrocyte glycophorin A. The K562 protein has an apparent MW identical to that of glycophorin A from red cells, it dimerises easily—a characteristic of the erythrocyte protein—it reacts with anti-glycophorin A antiserum, contains MN blood group activity, gives rise to CNBr fragments and contains glycopeptides/oligosaccharides which closely resemble those of glycophorin A (ref. 26). Furthermore, treatment of K562 cells with sodium butyrate induces erythroid differentiation including haemoglobin synthesis and the generation of erythrocyte-like particles²⁸.

A schematic model of the way in which we consider the biosynthesis of glycophorin A to occur is shown in Fig. 4. Three forms of glycophorin A can be recognised, GP_a, GP_b and GP_c. GP_c is the final glycophorin A molecule which is synthesised completely within the cell and is present at the cell surface after 35 min of pulse labelling. (This includes the 10 min required for trypsinisation.) This glycoprotein contains the terminal sialic acid residues and the penultimate galactosyl residues²⁶ and has

an apparent MW of 39,000 on 12% polyacrylamide gels. The presence of sialic acids gives it a higher electrophoretic mobility in the presence of SDS than after removal of these residues²⁷ (Fig. 2). GP_b, which can be obtained from GP_c only after neuraminidase treatment, has an apparent MW of 41,000.

GP_a is obviously a precursor of glycophorin A. It reacts with anti-glycophorin A antiserum and the precursor/product relationship is evident from the pulse-chase experiments (Fig. 1a). It has an apparent MW of 37,000, does not adsorb to *R. communis* lectin-Sepharose and does not contain sialic acids (Fig. 2B,C). The 4,000 MW difference between desialylated glycophorin A (GP_b) and GP_a should correspond to approximately 20 monosaccharides. This cannot be accounted for solely by the completion of the complex oligosaccharide at Asn 26 but must involve the addition of galactosyl groups to the 15 alkali-labile chains. Apparently, the sialic acids are then rapidly added because a precursor form corresponding to desialylated glycophorin A (GP_b) is never observed during the pulse-chase experiments. The addition of galactosyl and sialic acid residues is known to occur late in the biosynthesis of glycoproteins, mainly in Golgi membranes²⁹.

GP_a is found after 5 min of labelling in microsomes when the carbohydrate is protected from the action of trypsin (Fig. 3). Trypsin treatment decreases its apparent MW by 3,000 and the released peptide should correspond to residues 101(102)–131 of erythrocyte glycophorin A or tryptic peptide T₄ (ref. 19) located at the COOH-terminal end. When sealed, everted erythrocyte ghosts are treated with trypsin no reduction in the apparent MW of glycophorin A is observed³⁰. One possibility is that proteolytic cleavage does occur, but the 3,000 MW difference is not seen because glycophorin A gives a broad and diffuse band on electrophoresis of ghosts. Another, and more likely, possibility is that the COOH-terminal portion of newly synthesised glycophorin A in microsomes is more easily accessible to proteases than the molecule in the finished erythrocyte membrane where it may be covered by other proteins. Treatment of intact cells with trypsin releases much more from the NH₂-terminal end of the polypeptide in soluble form^{15,31}. This indicates that the COOH-terminal part already in the endoplasmic reticulum has a similar location in the red cell membrane.

The biosynthetic pathway of glycophorin A corresponds well to that described for the glycoprotein (G protein) of vesicular stomatitis virus. Glycophorin A has a more complex carbohydrate composition than the viral protein, involving two types of quite different oligosaccharides. The glycosylation of glycophorin A probably involves both lipid intermediates for the synthesis of the *N*-glycosidic oligosaccharide¹² and direct glyco-

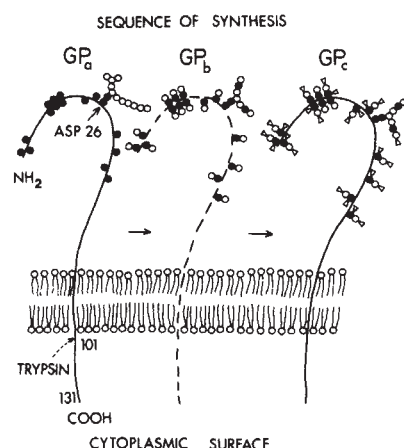


Fig. 4 Our schematic view of the biosynthesis of glycophorin A. GP_a, glycophorin A precursor; GP_b, glycophorin A treated with neuraminidase to demonstrate the stage before addition of sialic acid residues; GP_c, final glycophorin A. ● = hexosamine, ○ = neutral hexose, ▽ = sialic acid. The trypsin cleavage site of GP_a in microsomes is shown. The asparagine-linked oligosaccharide located at residue 26 in the polypeptide chain is shown in GP_a in an 'untrimmed' form from the structure described by Li *et al.*¹².

sylation of the serine/threonine residues from UDP-*N*-acetyl-D-galactosamine³². However, similar times are required for insertion into the endoplasmic reticulum membrane, glycosylation and appearance at the cell surface. Whether glycoporphin A contains an additional NH₂-terminal signal sequence found in most glycoprotein precursors³³ is not known and is being investigated.

We thank Anneli Asikainen, Marja Wilkman and Liisa Alajoki for assistance. K562 cells were obtained from Dr K. Nilsson. This study was supported by the Academy of Finland, the Finnish Cancer Society and Lääke Oy.

Received 12 February; accepted 8 May 1979.

1. Bretscher, M. S. & Raff, M. C. *Nature* **258**, 43–49 (1975).
2. Gahmberg, C. G. in *Dynamic Aspects of Cell Surface Organization* (eds Poste, G. & Nicolson, G. L.) 371–421 (North-Holland, Amsterdam, 1977).
3. Bretscher, M. S. *Nature new Biol.* **231**, 229–232 (1971).
4. Bretscher, M. S. *J. molec. Biol.* **98**, 831–833 (1975).
5. Nicolson, G. L. & Singer, S. J. *J. Cell Biol.* **60**, 236–248 (1974).
6. Gahmberg, C. G. & Hakomori, S. J. *J. biol. Chem.* **248**, 4311–4317 (1973).
7. Steck, T. L. & Dawson, G. J. *J. biol. Chem.* **249**, 2135–2142 (1974).
8. Rothman, J. E. & Lodish, H. F. *Nature* **269**, 775–780 (1977).
9. Katz, F. N., Rothman, J. E., Knipe, D. M. & Lodish, H. F. *J. supramolec. Struct.* **7**, 353–370 (1977).
10. Garoff, H., Simons, K. & Dobberstein, B. *J. molec. Biol.* **124**, 587–600 (1978).
11. Rothman, J. E., Katz, F. N. & Lodish, H. F. *Cell* **15**, 1447–1454 (1978).

12. Li, E., Tabas, I. & Kornfeld, S. *J. biol. Chem.* **253**, 7762–7770 (1978).
13. Steck, T. L. *J. Cell Biol.* **62**, 1–19 (1974).
14. Marchesi, V. T., Furthmayr, H. & Tomita, M. A. *Rev. Biochem.* **45**, 667–698 (1976).
15. Marchesi, V. T., Tillack, T. W., Jackson, R. L., Segrest, J. P. & Scott, R. E. *Proc. natn. Acad. Sci. U.S.A.* **69**, 1445–1449 (1972).
16. Furthmayr, H. *Nature* **271**, 519–524 (1978).
17. Segrest, J. P., Jackson, R. L., Andrews, E. P. & Marchesi, V. T. *Biochem. biophys. Res. Commun.* **44**, 390–395 (1971).
18. Tomita, M. & Marchesi, V. T. *Proc. natn. Acad. Sci. U.S.A.* **72**, 2964–2968 (1975).
19. Tomita, M., Furthmayr, H. & Marchesi, V. T. *Biochemistry* **17**, 4756–4770 (1978).
20. Gahmberg, C. G., Jokinen, M. & Andersson, L. C. *Blood* **52**, 379–387 (1978).
21. Tanner, M. J. A. & Anstee, D. J. *Biochem. J.* **155**, 701–703 (1976).
22. Dahr, W., Uhlenbruck, G., Leikola, J., Wagstaff, W. & Landfield, K. *J. Immunogenet.* **3**, 329–346 (1976).
23. Gahmberg, C. G., Myllylä, G., Leikola, J., Pirkola, A. & Nordling, S. *J. biol. Chem.* **251**, 6108–6116 (1976).
24. Lozzio, C. B. & Lozzio, B. B. *Blood* **45**, 321–334 (1975).
25. Andersson, L. C., Nilsson, K. & Gahmberg, C. G. *Int. J. Cancer* **23**, 143–147 (1979).
26. Gahmberg, C. G., Jokinen, M. & Andersson, L. C. *J. biol. Chem.* (in the press).
27. Gahmberg, C. G. & Andersson, L. C. *J. biol. Chem.* **252**, 5888–5894 (1977).
28. Andersson, L. C., Jokinen, M. & Gahmberg, C. G. *Nature* **278**, 364–365 (1979).
29. Schachter, H. *et al. J. biol. Chem.* **245**, 1090–1100 (1970).
30. Steck, T. L. in *Membrane Research* (ed. Fox, C. F.) (Academic, New York, 1972).
31. Gahmberg, C. G. *J. biol. Chem.* **251**, 510–515 (1976).
32. McGuire, E. J. & Roseman, S. *J. biol. Chem.* **242**, 3745–3747 (1967).
33. Blobel, G. & Dobberstein, B. *J. Cell Biol.* **67**, 835–851, 852–867 (1975).
34. Laemmli, U. K. *Nature* **227**, 680–685 (1970).
35. Bonner, W. M. & Laskey, R. A. *Eur. J. Biochem.* **46**, 83–88 (1974).
36. Rice, R. H. & Means, G. E. *J. biol. Chem.* **246**, 831–832 (1971).
37. Gahmberg, C. G. & Andersson, L. C. *J. exp. Med.* **148**, 507–521 (1978).
38. Sage, J. H. & Green, R. W. *Meth. Enzym.* **28**, 332–339 (1972).

The three-dimensional structure of tubulin protofilaments

L. A. Amos & T. S. Baker*

MRC Laboratory of Molecular Biology, Cambridge, UK

A three-dimensional image of tubulin protofilaments, reconstructed to 20 Å resolution from electron micrographs of negatively stained zinc-induced sheets, has been used to generate an improved model for microtubule substructure. This model has been used to phase a reconstructed image from X-ray amplitudes, which is compatible with the electron microscope results.

THE three-dimensional structure of eukaryotic microtubules has been studied by electron microscopy (EM) of individual specimens^{1–4} and by X-ray diffraction of oriented gels of microtubules^{5–7}. Results from EM suggested an arrangement of monomer subunits on the same helical surface lattice for both neurotubules and flagellar microtubules^{2,3}. Early X-ray diffraction patterns seemed to be in disagreement⁵, but this was resolved after further studies^{6,7}; the X-ray data now seem to be quite consistent with the geometry determined by EM. However, there still seem to be some differences in the data, especially with respect to the shapes of the subunits, even at very low (40 Å) resolution. Even different EM studies have tended to give somewhat different impressions of the substructure. In images of brain tubulin sheets, which are thought to be the equivalent of opened-out microtubules, each tubulin subunit seemed to be split longitudinally into two separate domains^{2,4}. This feature was not detected in a three-dimensional analysis of intact flagellar microtubules³, in which the monomers apparently consisted of single globular units, although the overall outline of the subunits was similar. Furthermore, a model based on an interpretation of the intensity distribution in X-ray diffraction patterns of brain microtubules⁷ showed

subunit shapes which were incompatible with either of the classes of image obtained by EM. We now present a new model of the structure of tubulin, based on three-dimensional EM image reconstruction of extended brain tubulin sheets, which seems to explain the differences between the earlier images and suggests a compatible interpretation of the X-ray diffraction patterns from microtubules.

Zinc-induced tubulin sheets

The specimens we studied were the large two-dimensional crystalline arrays of brain tubulin which assemble in the presence of zinc ions^{8,9}. These sheets are apparently made up of the same linear arrays of subunits, or protofilaments, as 'normal' sheets, which are often found in samples of repolymerised microtubule protein and which have been proposed as precursors in microtubule assembly^{2,10}. Whereas normal sheets are usually no more than 12 protofilaments wide, presumably because a sheet of 13 protofilaments closes to form a microtubule, zinc-induced sheets may be 100 or more protofilaments wide and are therefore much better for periodic structural analysis. Studies of their projected structure have shown that protofilaments are arranged with alternating polarity (see Fig. 1a), so that these arrays have P₂₁ crystallographic symmetry^{11,12}. The arrangement is thus quite different from that in a microtubule or a normal sheet, where all the protofilaments have the same polarity. However, the individual protofilaments seem to be very similar, having the same 48–49 Å side-to-side spacing in both types of sheet. This common feature suggests that the rotational orientation of a protofilament about its axis is similar in both structures. By comparing the obliquity of the projected subunit shapes in Fig. 1a with those in computer-filtered EM images of normal tubulin sheets^{2,4}, we were also able to deduce that the protofilament images labelled I in Fig. 1 represent the view from the outside of a normal microtubule¹¹.

* Present address: Rosenstiel Basic Medical Sciences Research Center, Brandeis University, Waltham, Massachusetts 02154.

Figure 1a was derived from four highly selected low-dose EM images of negatively stained specimens and includes data to 16 Å resolution¹¹. To obtain a three-dimensional image, this untitled projection was combined with 40 independent images of sheets tilted through angles of 21°, 35° or 51° about various axes¹⁴. The number of different orientations of the sheets was more than sufficient to provide homogeneous data to 20 Å resolution within the angular range covered by tilting up to 51°. The intensities of most of the reflections fell to noise level within this range. Additional data, corresponding to higher tilt angles, which could not easily be reached by tilting in the microscope, were obtained from images of negatively stained specimens embedded in plastic and sectioned normal to the sheets¹³. These data were to about 40 Å resolution; the complete set of data is therefore still not fully isotropic, so that density variations in the direction normal to the sheets are not quite as well resolved as those in the plane of the sheets. Full details of data collection and analysis are presented elsewhere¹³.

Three-dimensional model of the sheets

The main features of the reconstructed three-dimensional density distribution are illustrated in Figs 1b and 2. Viewed down its long axis, each protofilament is roughly triangular in shape, with one corner of the triangle protruding from one side of the sheet. Thus, the two exposed surfaces of a protofilament are very different. In longitudinal view, one surface has the form of a longitudinal ridge and the other has a flat ladder-like appearance, with sloping bi-lobed rungs. If we are correct in assuming that the protofilament images labelled I in Fig. 1a represent the view from outside a microtubule, then the ridged surface must be on the outside, while the ladder-like appearance should occur on the inside surface. This is discussed in detail later.

The apparent thickness of the negatively stained sheets, as observed directly in section and also as estimated from the tilted views, is only about 45–50 Å, compared with a difference of 70–80 Å between the inner and outer radii of a microtubule, as estimated from X-ray diffraction measurements^{5–7}. Therefore, the sheets have probably shrunk in this dimension as they dried down on to the EM support grid. It is also possible that part of the apparent reduction in thickness results from the masking of

fine surface features extending into the negative stain, or by positive staining of protein on the protofilament surface. Thus, the ridge in particular may extend out further than it seems to in the reconstructed image. The importance of such effects is discussed later.

As described previously¹¹, each protofilament clearly consists of an alternating linear array of two kinds of slightly different subunits, believed to represent the two species of 55,000 molecular weight monomer which form the tubulin heterodimer¹⁵. Both kinds of subunit appear in projection as elongated shapes (Fig. 1a), with their long axes at about 40° to the protofilament axes. If one attempts to follow the density corresponding to an individual subunit through successive sections parallel to the central plane of the sheet (Fig. 2), the peaks move gradually downwards between the first and last sections. The subunits thus apparently have the general form of tilted ellipsoids when viewed from the side, as well as face on.

The difference between the two kinds of subunit seems to be rather less in three dimensions than in the two-dimensional projection, showing significantly only at the ladder-like inside surface of the protofilaments. The apparently more definite division into two domains of one of the two subunits presumably represents a real difference in the internal structure of the two monomers.

The pattern of interprotofilament contacts is also not as complex as suggested by the original projection (Fig. 1a), where the contacts between each pair of subunits seem to be represented by two separate regions of higher density in the interprotofilament gap¹¹. In the three-dimensional model, the contacts seem to take place over single fairly broad zones (Fig. 2, P to Q or R to S), centred on the central plane of the sheets. The complexity in projection is due to superposition of the intricate stain distributions at the two surfaces of the sheets.

Model protofilaments rearranged to represent a microtubule

Normal microtubules observed *in situ* consist, in most cells, of 13 longitudinal protofilaments¹⁶ in a polar arrangement³. Reassembled microtubules often contain 14 protofilaments¹⁷, but these are probably aberrant structures. To simulate a normal

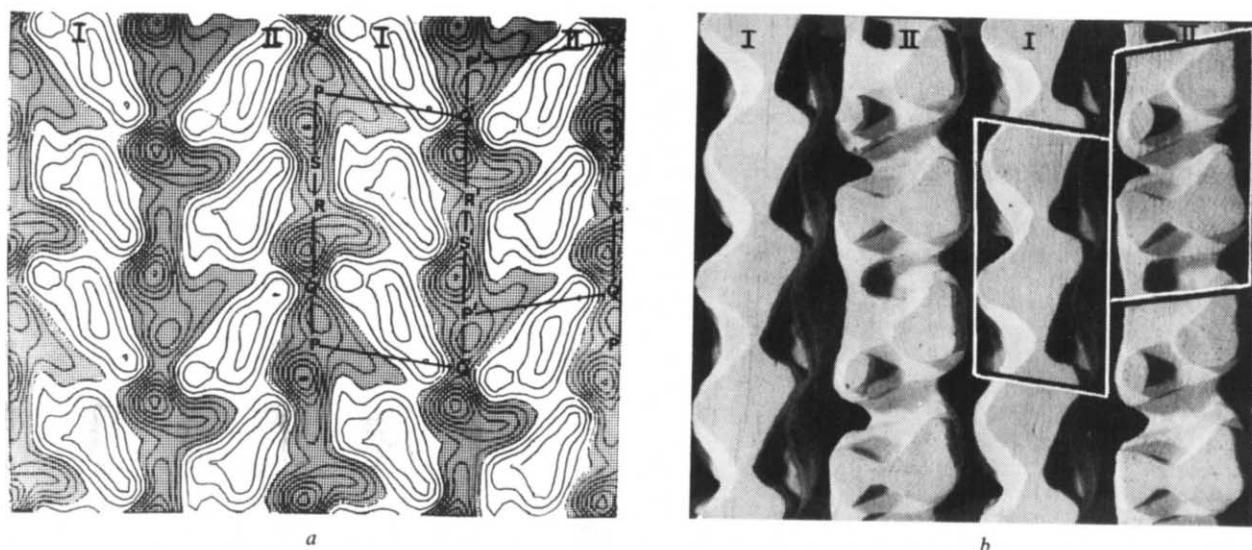


Fig. 1 a, Computer reconstructed image of negatively stained zinc-tubulin sheets seen in projection¹¹. The variation in density is represented by evenly spaced contour levels. Regions dominated by negative stain are shown shaded while absence of stain, corresponding to protein, is white. Neighbouring longitudinal protofilaments, each 48 Å wide, are related by P2₁ symmetry, about screw dyad axes normal to the protofilaments and in the plane of the sheet. Each globular unit along a protofilament is thought to represent a monomer of α - or β -tubulin. The boxes drawn on the image each enclose what seems to be an 80 Å long α - β heterodimer. b, A solid model of part of a zinc-tubulin sheet, with boundaries which follow one of the contour levels in Fig. 2. Two different surfaces of the protofilaments are presented, as adjacent protofilaments face different ways.

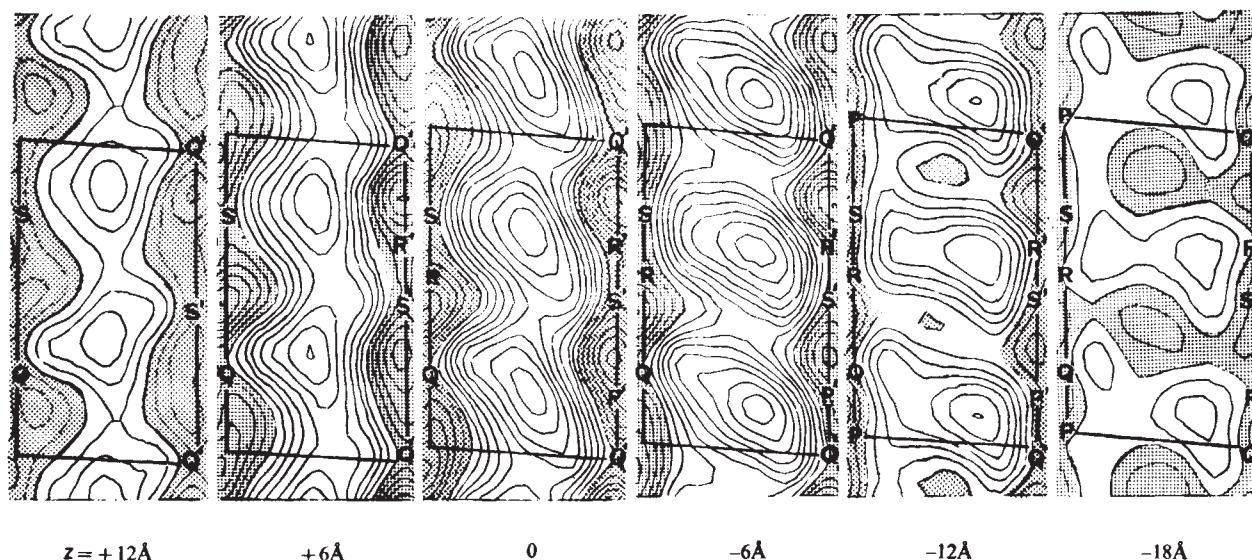
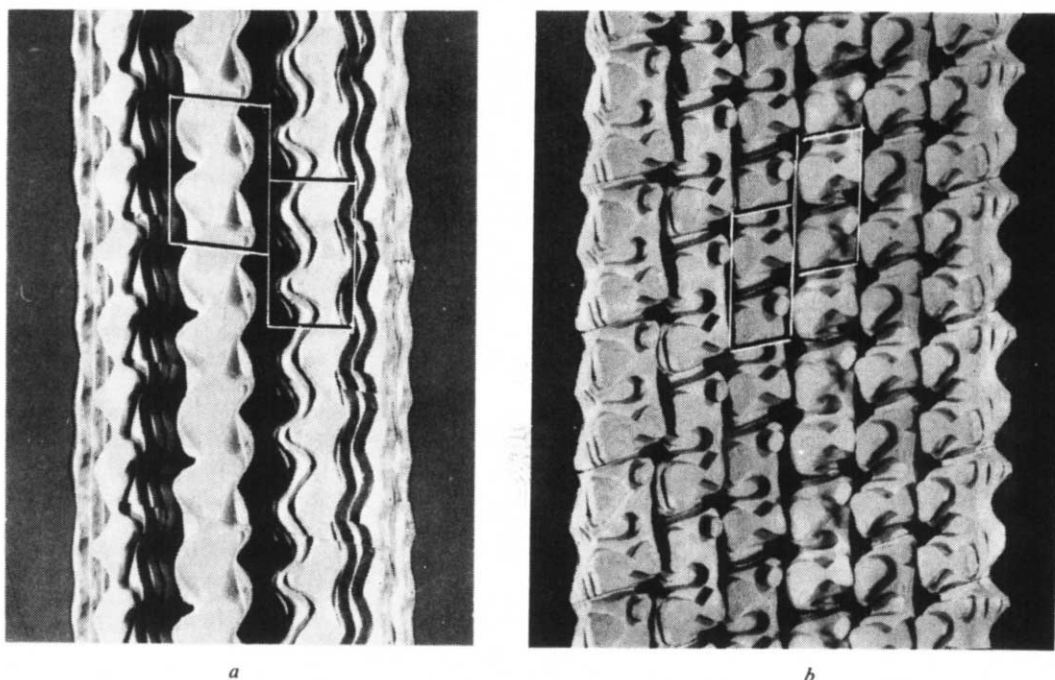


Fig. 2 Sections at different levels (z) through one protofilament of the reconstructed three-dimensional model of the negatively stained zinc-tubulin sheets. Below each section is given its distance from the plane through the centre of the sheet. Features at the boundaries of the protofilament are labelled as in the projection in Fig. 1a.

microtubule, therefore, we have arranged 13 model zinc-tubulin protofilaments, as shown in Fig. 3. In this arrangement, the former regions of contact between subunits in the sheet seem to match up again in the tubule, although in different polar orientations¹¹. Probably the same hydrophobic regions of each subunit are involved in the assembly of both types of structure. The model in Fig. 3 was assembled with the ridged sides of the protofilaments outermost, so that the projected subunit shapes would agree with images of negatively stained normal sheets^{2,4}. There are several observations which support this choice of orientation. One is the good agreement between the appearance of this model and that of the three-dimensional model reconstructed from images of intact flagellar microtubules³. In particular, the deep interprotofilament grooves are features of the

outer surfaces of both models. The presence of these grooves at high radius is also strongly indicated by X-ray diffraction data⁵⁻⁷. If the wedge-shaped protofilaments were oriented with the ridged surface inwards, grooves would not be formed specifically at the outer surface; the interprotofilament contrast would be weaker and roughly equal at all radii between 70 and 150 Å. Clear images of sectioned normal microtubules^{17,18} also support the conclusion that the protofilaments should appear in cross-section as outwardly pointing wedge shapes. A final observation which indicates that, in a normal microtubule, interprotofilament contact is limited to inner radii (less than 110 Å) is that the spacing of protofilaments in flat sheets (normal or zinc-induced) is consistently less (48–49 Å) than the spacing of 53 Å at the mean radius (110 Å) of a microtubule.

Fig. 3 *a*, Models of zinc-tubulin protofilaments from the reconstructed image shown in Fig. 1b, rearranged to conform to the helical symmetry established for a normal microtubule. Thus, the monomer subunits lie on a family of three shallow left-handed helices. The relative arrangement of tubulin dimers is that observed in complete flagellar microtubules³ (see text). Two unit cells, each enclosing an 80-Å long α - β dimer, are outlined. *b*, Inside surface of part of a microtubule, opened out to form a 'normal' sheet. The polar arrangement of the protofilaments is the same as in *a*. Comparison of these images with the reconstructed three-dimensional model of intact flagellar microtubules³, suggests that the bottom of the model, as shown in these two views, corresponds to the basal end of a flagellar microtubule.



The other assumption incorporated in the model shown in Fig. 3 is that the subunit dimers in adjacent protofilaments are in a staggered arrangement, as observed in complete flagellar microtubules, that is, the A-tubules of outer doublet microtubules³ and the central singlet microtubules¹. This is in fact the only helically symmetrical arrangement possible for a 13-protofilament microtubule in which the monomer subunits are arranged as shown. There is some evidence^{7,12} that brain tubulin may sometimes reassemble *in vitro* in an alternative arrangement similar to that in a flagellar B-tubule³, but when microtubules are reassembled from brain extracts containing a full complement of so-called microtubule associated proteins, the latter, seen as projections from the outsides of the tubules, seem to follow the symmetry of the A-tubule lattice¹⁹. This point is discussed in more detail elsewhere²⁰.

Variations in EM images

It is now understandable why somewhat variable images have been obtained in previous EM studies of protofilament structure. The appearance seems to vary according to which of the two different protofilament surfaces is the more strongly

emphasised by the negative stain. The bi-lobed, or ladder-like appearance obtained for opened-out brain microtubules^{2,3} corresponds rather closely to the appearance of the inner surface of the present model. Erickson² noted that the normal sheets, being curved in solution, always fell with the outer surface next to the carbon substrate of the EM grid. Presumably the inside surface, which was thereby most exposed to the negative stain subsequently applied to the grid, was more strongly contrasted than the outer surface. The outer surface of the three-dimensional image reconstructed from images of intact flagellar microtubules³, on the other hand, closely resembles the outer surfaces of the protofilaments in the present model, whereas the inner surface showed much less detail. In the case of an intact microtubule, it seems that the outside is most strongly contrasted by negative stain. In the zinc-induced sheets the orientation of the protofilaments alternates. Differential staining of the two sides of such sheets probably explains the great difference in projection of adjacent protofilaments observed initially by Crepeau *et al.*⁴. However, the images we have used all show very good P2₁ symmetry¹¹, so we deduce that in our specimens each protofilament must be more or less equally well contrasted on both sides. The three-dimensional image reconstructed from

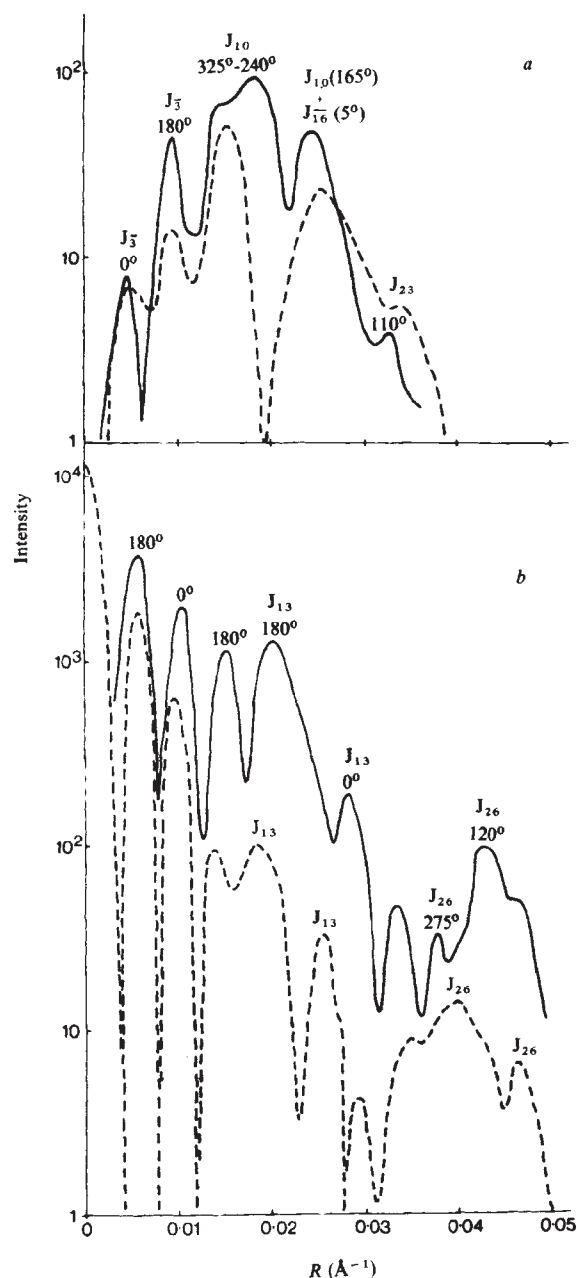


Fig. 4 Intensity distributions for the equatorial (*b*) and 40 Å (*a*) layer lines of diffraction patterns from real and model microtubules. The dashed lines represent the intensity distributions measured by Mandelkow *et al.*⁷ in X-ray diffraction patterns obtained from oriented hydrated gels of microtubules. The solid lines follow cylindrically averaged intensity distributions calculated for a model microtubule, consisting of 13 zinc-tubulin protofilament images arranged on a mean radius of 110 Å; the density distributions were uniformly 'stretched' by 30% in the direction normal to the zinc-tubulin sheets, to correspond more closely to the microtubule dimensions estimated from X-ray diffraction data⁷ (see text). The only significant effect of this is a slight change in the radial positions of the intensity peaks in the calculated diffraction pattern. The relative phases (for protein represented as positive density) determined for the various helical contributions are indicated above the intensity peaks. On the calculated equator, the strong peaks near the origin represent the transform of a hollow cylinder. The peaks at 0.020 Å⁻¹ and 0.028 Å⁻¹ (labelled J₁₃) arise mainly from the division of the cylinder into 13 distinct longitudinal protofilaments: the first is from 13-fold contrast at the outside of the cylinder, the second from the much weaker contrast on the inside. Second order diffraction from the protofilaments provides the main contributions to the last group of peaks (J₂₆): the weak inner peak of this group corresponds to features of the outside of the cylinder, the stronger peak to the inner side. On the 40 Å or second layer line (based on an axial repeat of 80 Å) of the calculated diffraction pattern, the two (J₃) peaks close to the meridian arise from outer and inner radii contributions from the family of three shallow left-handed helices; the apparently double third peak (J₁₀) (with varying phase across it) consists mainly of contributions from the 10 steeper, right handed helices at outer radii. The broad fourth peak includes a small contribution from the latter family at inner radius, but also significant contributions from the left-handed 16-member family at radii centred on a mean of about 110 Å. The calculated curves do not correspond exactly with those measured from X-ray diffraction patterns; for example, the relative intensities on the equator fall off differently and the J₁₀ peak on the 40 Å layer line has a different shape. However, some differences are to be expected, as the model was derived by studying dehydrated protein embedded in metal stain, whereas the X-ray diffraction pattern is from hydrated unstained specimens and provides information about the contrast between protein and aqueous solvent.

them should therefore provide a more reliable impression of the structure than has been obtained from previous EM studies of tubulin structure.

However, because of artefacts due to dehydration, radiation damage and the limitations of negative stain contrast, the results obtained by EM are not wholly satisfactory for studying the native structure of proteins. Low-dose electron microscopy of unstained specimens, embedded in glucose or sucrose solutions to alleviate the effects of dehydration, has proved successful in studies of the substructure of certain thin two-dimensional crystals of biological origin^{14,21}. However, the results obtained so far with tubulin sheets⁴ have not been very promising, the resolution attained being somewhat less than that obtained from negatively stained specimens. The zinc-tubulin sheets, although they grow much larger than normal tubulin sheets, probably still do not allow sufficient averaging to compensate for the extremely low signal-to-noise ratio in EM images of unstained protein.

Diffraction patterns from microtubules

X-ray diffraction of oriented gels of hydrated microtubules provides information about the protein in its native state, but the X-ray patterns from native specimens do not provide sufficient data by themselves for the reconstruction of a three-dimensional image of the structure. One way of trying to supply the missing information at low resolution is to use phase information obtained by electron microscopy⁷. For this purpose, we have calculated a three-dimensional diffraction pattern for the model density distribution shown in Fig. 3. The calculated diffracted intensity distributions, averaged around the vertical axis to simulate diffraction from an oriented gel of microtubules, are shown in Fig. 4 for the equatorial and 40 Å layer lines, together with the corresponding relative intensities measured from X-ray diffraction patterns. These curves calculated for the rearranged zinc-tubulin protofilaments are surprisingly more similar to the X-ray curves than are the distributions obtained directly from optical diffraction patterns of negatively stained microtubules^{1-4,7}. In the latter case, the peak closest to the meridian on the 40 Å layer line is usually the strongest on that

layer line, a peak corresponding to the third peak on the X-ray 40 Å layer line is much weaker in the microtubule optical diffraction patterns and the other two X-ray peaks are absent. As discussed above, this seems to be because negative stain does not map out the surface of a protofilament in an intact microtubule as closely as it does that of a protofilament in a zinc-tubulin sheet.

Interpreting the X-ray patterns

Cohen *et al.*^{6,7} have interpreted the low angle X-ray intensity distribution (Fig. 4, dashed lines) in terms of diffraction mainly from the inner and outer surfaces of the microtubules, which apparently have inner and outer radii of 70 and 150 Å respectively. The strong equatorial J_{13} and J_{26} reflections have both been identified as coming from the outer surface, with mean radii of 126 and 115 Å respectively. The series of four peaks on the 40 Å layer line were indexed as two J_{-3} contributions from outer and inner mean radii of 130 and 70 Å, followed by two J_{10} contributions from outer and inner mean radii of 120 and 70 Å.

Mandelkow *et al.*⁷ attempted to reconstruct a three-dimensional model of a microtubule from the measured X-ray intensities, interpreted in the above manner, by combining them with phases determined by Erickson², and confirmed by their own observations, from two-dimensional EM images of microtubules opened out as sheets. (In calculated diffraction patterns from digitized images, with a phase origin equivalent to that used in Fig. 4, phases of 180° are found for each of the lattice points corresponding to single J_{13} , J_{26} , J_{-3} and J_{10} contributions.) However, as the relative intensities in the X-ray diffraction pattern and the diffraction patterns from the original EM images were very different, the hybrid reconstructed image⁷ was also quite different from the original images, with elongated subunits sloping in the opposite direction. Without reasonable agreement between the two techniques, there is no justification for combining EM phases with X-ray amplitudes. It is now possible, however, from our analysis of the zinc-tubulin structure, to suggest some modifications to the above interpretation and phasing of the X-ray patterns, which lead to a much closer consistency between the X-ray and EM data.

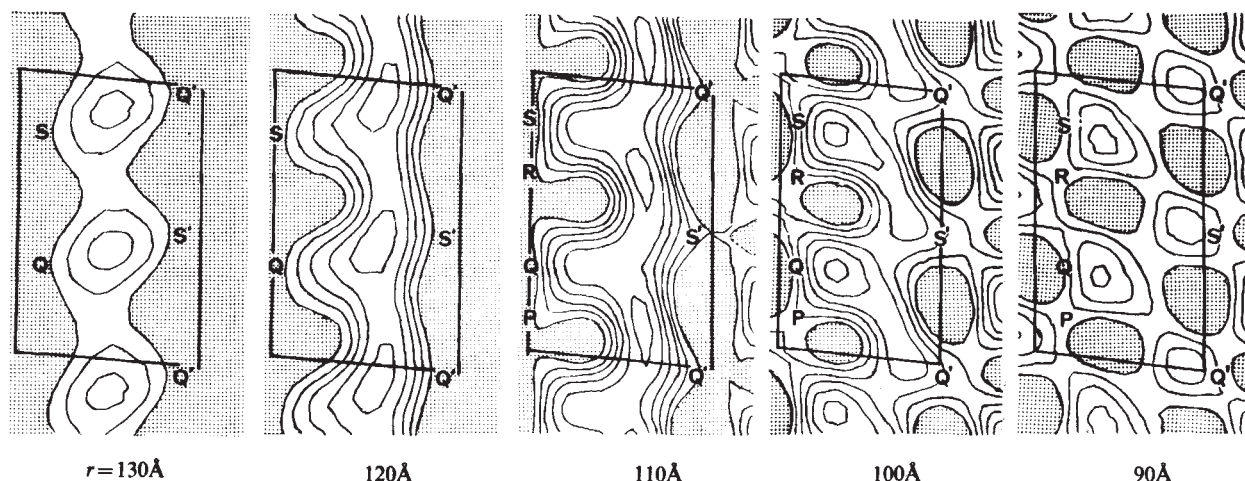


Fig. 5 Sections at different radii (r) through part of the reconstructed density distribution, calculated using Fourier amplitudes deduced from X-ray diffraction patterns of intact brain microtubules and phases (see Fig. 4) derived from model zinc-tubulin protofilaments rearranged as shown in Fig. 3. Features of the protofilament outline, which seem to correspond to those in the sections through the original image (Fig. 2), are labelled P to S. The two images are rather different at smaller radii; the sloping regions of high density (S to S') and (Q to Q') have a steeper slope than the corresponding regions in Fig. 2. (This difference seems to be less for one type of subunit in Fig. 2 than for the other.) As well as the equatorial and second (40 Å) layer lines shown in Fig. 4, the data included weak contributions from the third and fourth layer lines, estimated from the published X-ray patterns⁷. (Note that the indexing of the odd layer lines is determined by the arrangement of 80 Å long tubulin dimers, which is discussed below Fig. 3.) On the third layer line, an apparent J_2 contribution to the X-ray diffraction patterns was included with a phase of 45° but being weak has very little effect. The fourth layer line data consisted of approximately equal J_{-6} and J_7 contributions with phases of 210° and 50°, based on the model calculations.

The main difference from the previous interpretation is the source of the fourth peak from the meridian on the 40 Å layer line. Calculations from our model suggest it is probably a mixture of different contributions. As well as a small J_{10} contribution, it is likely to include a major J_{-16} contribution, with a mean radius of about 110 Å. It is not completely clear why a J_{-16} contribution is never observed in optical diffraction patterns from negatively stained microtubules or normal tubulin sheets, but it is probably for the same reason that the J_{10} contribution is weaker than the J_{-3} contribution, instead of stronger as in the X-ray patterns. The J_{-3} reflections clearly do come from the innermost and outermost surfaces of the structure and must be relatively easily revealed by negative stain, provided the surface in question is exposed to it. On the other hand, according to the present interpretation of the X-ray data, both the major J_{10} and J_{-16} contributions come from radii near the middle of the protein layer, into which it is presumably more difficult for the stain to penetrate. It may be relevant in this context that the X-ray diffraction patterns from dehydrated microtubules are said to be more similar to the optical diffraction patterns from electron micrographs^{5,6}. This suggests that dehydration may cause the cleft between protofilaments to close up at inner radii and prevent features at these radii from contributing to the diffraction patterns.

Phasing the X-ray data

The relative phases we assign to the various contributions (Fig. 4), which were derived from the model shown in Fig. 3, also differ from those used by Mandelkow *et al.*⁷. This is mainly because, in the images of the flat sheets which supplied their phases, contributions from different radii in the microtubules become superimposed in projection. Unless features at all radii have corresponding spatial frequencies exactly in phase, the relative phases in the transform of such a flattened sheet will not agree with a separated set of either inside or outside surface contributions.

One significant difference is in the values assigned to the equatorial J_{26} intensity peaks. In EM images, the feature giving rise to the strongest diffraction contribution of this order is the shallow longitudinal groove bisecting one surface of each protofilament. We interpret this surface as being on the inside of a microtubule. In the X-ray diffraction patterns, the strongest J_{26} contribution seems to come from a radius of about 115 Å, which is within the extent of the deep interprotofilament grooves on the outside surface of the microtubule. Mandelkow *et al.*⁷ originally related this contribution to the apparent splitting of the protofilaments (hence the 180° relative phase they assigned to it), but we now interpret it as most likely due to the sharpness of the ridges on the outer surfaces of the protofilaments. A smaller peak further out on the equator, at $R = 0.047 \text{ Å}^{-1}$, probably arises from the apparent splitting of the protofilaments at inner radii.

According to this interpretation, the relative intensities of the two J_{26} contributions, from the inner and outer surfaces of the microtubules, are reversed in the X-ray diffraction patterns and the patterns calculated from the rearranged zinc-tubulin model. This could be due to artefacts of negative staining whereby the tips of the ridges on the outer surface of each protofilament might be swamped by negative stain and the outer surface J_{26} contribution thus reduced, while the shallow groove down the centre of the inside surface of the protofilament would tend to be over-emphasised by the stain, causing the inner surface J_{26} contribution to be enhanced.

Model from combined X-ray and EM data

A test of the above interpretation of the X-ray patterns is to see whether a reconstructed image, based on the X-ray amplitudes and the phases derived from the rearranged model, resembles the results from EM alone. Figure 5 shows sections through one protofilament of such a calculated microtubule density distribution. This structure can be compared with the zinc-tubulin

image shown in Fig. 2, bearing in mind that the latter structure seems to have shrunk by 30–40% in the z -direction. There are some other differences between the two structures, as expected from the differences in the two sets of diffraction curves in Fig. 4. The sections in the region of the outside surface (around $z = 12$, 6 Å in Fig. 2; around $r = 130$, 120 Å in Fig. 5) are reasonably similar, but at inner radii there are some distinct differences. In particular, the 'rungs of the ladder' (around $z = -18$, -12 Å in Fig. 2; around $r = 90$, 100 Å in Fig. 5) seem much more steeply tilted, because of the greater strength of the J_{-16} contribution to the data, relative to the J_{-3} contribution. Also, the differences between the two species of tubulin subunit are much less obvious, because of the relative weakness of the third (27 Å) layer line in the X-ray diffraction pattern (no data from other odd-numbered layer lines were included).

We do not know the reason for these remaining differences. Discrepancies in the central sections of the protofilaments could obviously be due to failure of the negative stain to penetrate fully into this region, even in the case of the zinc-tubulin sheets, which seem to stain better than normal tubulin structures. But this explanation is not adequate for the differences observed in the surface structure. One must postulate that the negative stain does not adequately represent the native protein contours. This might be due to selective positive staining²² of certain parts of the protein molecules. Alternatively, there may be reproducible changes in conformation during negative staining and dehydration of the EM specimens, as there seems to have been a consistent shrinkage of the sheets. Neither of these possible effects could be corrected for when the negatively stained density distribution was 'stretched' (Fig. 4) to adjust the dimensions to those of hydrated protein. Because of this lack of exact correspondence between the two images, we still cannot be sure that the relative phases, especially those related to features on the inside of the cylinder, are quite correct.

Resolution of the remaining uncertainties in the structure at this resolution must await an independent determination of the Bessel function assignments and phases in the X-ray diffraction pattern, for example, by labelling the protein with heavy atoms²³, if this proves possible. However, the consistency between the results obtained by electron microscopy and X-ray diffraction is clearly much better than it originally seemed to be⁵⁻⁷ and, for this reason, we believe that the three-dimensional model of zinc-tubulin protofilaments presented here is a good first approximation to the low resolution structure of assembled tubulin.

We thank Drs A. Klug, R. Henderson and P. N. T. Unwin for advice. This work was aided by a grant from the Jane Coffin Childs Memorial Fund for Medical Research. During the course of the work, T.S.B. was a fellow of the Jane Coffin Childs Memorial Fund.

Received 16 February; accepted 8 May 1979.

- Chasey, D. *Expl Cell Res.* **74**, 140–146 (1972).
- Erickson, H. P. *J. Cell Biol.* **60**, 153–167 (1974).
- Amos, L. A. & Klug, A. *J. Cell Sci.* **14**, 523–549 (1974).
- Crepeau, R. H., McEwen, B., Dykes, G. & Edelstein, S. J. *J. molec. Biol.* **116**, 301–315 (1977).
- Cohen, C., Harrison, S. C. & Stephens, R. E. *J. molec. Biol.* **59**, 375–380 (1971).
- Cohen, C., DeRosier, D. J., Harrison, S. C., Stephens, R. E. & Thomas, J. *Ann. N.Y. Acad. Sci.* **253**, 53–59 (1975).
- Mandelkow, E., Thomas, J. & Cohen, C. *Proc. natn. Acad. Sci. U.S.A.* **74**, 3370–3374 (1977).
- Larsson, H., Wallin, M. & Edstrom, A. *Expl Cell Res.* **100**, 104–110 (1976).
- Gaskin, F. & Kress, Y. *J. biol. Chem.* **252**, 6918–6924 (1977).
- Kirschner, M. W., Williams, R. C., Weingarten, M. & Gerhart, J. C. *Proc. natn. Acad. Sci. U.S.A.* **71**, 1159–1163 (1974).
- Baker, T. S. & Amos, L. A. *J. molec. Biol.* **123**, 89–106 (1978).
- Crepeau, R. H., McEwen, B. & Edelstein, S. J. *Proc. natn. Acad. Sci. U.S.A.* **75**, 5006–5010 (1978).
- Amos, L. A. & Baker, T. S. *Int. J. biol. Macromolec.* (submitted).
- Henderson, R. & Unwin, P. N. T. *Nature* **257**, 28–32 (1975).
- Ludueña, R. F., Shooter, E. M. & Wilson, L. *J. biol. Chem.* **252**, 7006–7014 (1977).
- Tilney, L. G. *et al. J. Cell Biol.* **59**, 267–275 (1973).
- Pierson, G. B., Burton, P. R. & Himes, R. H. *J. Cell Biol.* **76**, 223–228 (1978).
- Warner, F. D. & Satir, P. *J. Cell Sci.* **12**, 313–326 (1974).
- Amos, L. A. *J. Cell Biol.* **72**, 642–654 (1977).
- Amos, L. A. in *Microtubules* (eds Roberts, K. & Hyams, J.S.) (Academic, London, 1979).
- Unwin, P. N. T. & Henderson, R. *J. molec. Biol.* **94**, 425–440 (1975).
- Unwin, P. N. T. *J. molec. Biol.* **98**, 235–242 (1975).
- Holmes, K. C., Stubbs, G. J., Mandelkow, E. & Gallwitz, U. *Nature* **254**, 192–196 (1975).

letters

Polar clearing in the Venus clouds observed from the Pioneer Orbiter

PIONEER Venus 1 was put into a 24-h orbit around the planet on 4 December 1978. Since then, it has made remote sensing observations of the clouds and the overlying atmosphere at IR, visible and UV wavelengths. The IR instrument includes a channel at a wavelength of $11.5\ \mu\text{m}$, which is used to measure the effective temperature of the cloud tops. Carbon dioxide, and all of the known constituents of the gaseous atmosphere, are highly transparent at this wavelength. Earth-based telescopic observations at similar wavelengths¹⁻³ have generally shown fairly uniform temperatures of the order of 240 K across the face of Venus. The accepted interpretation has been that the cloud cover on Venus not only covers the whole planet, but is also extremely uniform. The early Pioneer observations confirmed this general picture, for the equatorial and mid-latitudes which comprise most of the surface area of the planet⁴. Interesting and meteorologically important structure with diurnal, seasonal and random components is found in the Pioneer measurements, as it was in earlier ground-based and spacecraft observations, but this structure has an amplitude of 10 K or less¹⁻⁶. However, these subtle contrasts which dominate thermal maps of lower latitudes change to dramatic structure near the pole. Even the Earth-based observers, who must view the polar regions on Venus at very oblique angles, have reported cold bands surrounding one or other of the poles at various times² and occasional polar 'hot spots'. The Pioneer orbit passes almost directly over the pole (inclination = 105°) and so provides the first good views of this interesting region. The IR instrument was

operated only in its low spatial resolution mode for the first part of the mission. Results from those observations⁴ showed a wave-shaped collar of high, cold cloud surrounding the north pole, with warmer cloud temperatures polewards. Close to the pole itself, the observed temperature is the highest anywhere on the planet, not only at the cloud top but also in the overlying atmosphere⁴. We report here the first high data rate, high spatial resolution observations of the polar region. These were obtained on the tenth orbit of Venus on 15 December 1978.

Figure 1 shows a single swath across the planet, covering latitudes from about 40° to 80°N as shown in Fig. 2. Figure 3 shows an image of the region from a series of contiguous swaths. The individual swaths are obtained as a result of the spinning motion of the spacecraft, which rotates at $\sim 5\text{ r.p.m.}$ The separation of the swaths is due to the motion of the spacecraft along its orbit.

These data show an extremely localised, very warm feature at about 80°N and 359°E (corresponding to a local time of day of about 9 am). The peak temperature of 260 K is the highest cloud-top temperature ever to be observed on Venus. Because of the abrupt nature of the increase, and also because the Pioneer IR radiometer makes independent measurements of the temperature of the clear atmosphere above the clouds, it is evident that the large temperature contrast is due to a depression or clearing of the cloud layer rather than to an actual temperature increase at a constant height level. This in turn is probably due to strong downwelling in the atmosphere near the pole. To account for the contrast of approximately 45 K between the peak of the anomaly and the nearby cold band which surrounds it, a difference of cloud height of approximately 15 km is required. This follows because the measured temperature lapse rate in the polar region is close to 3 K km^{-1} (ref. 5). Cloud height differentials of this magnitude are seldom, if ever, observed on the Earth.

It is interesting to consider how high the temperature would have to rise in an anomaly of this kind in order to correspond to a completely cloud-free atmosphere. To investigate this, we have calculated the transmission profile of a model Venus atmosphere corresponding to the spectral bandpass of the $11.5\ \mu\text{m}$ channel. As stated previously, this wavelength was selected because it is one of the most transparent in the thermal IR. However, because of the high concentration of CO_2 on Venus, even this would become opaque (defined as an optical depth greater than unity) by a depth corresponding to a pressure of about 0.8 bar. At this level, the atmospheric temperature is approximately 290 K. Thus, the highest temperature measured approaches, but definitely does not attain, that which would correspond to a totally cloud-free region. However, as can be seen from Fig. 3, the high resolution data sequence ended before the coverage of the feature was completed. The temperature may have risen higher, perhaps even to the temperature corresponding to a total absence of cloud, nearer to the pole itself. An analysis of further high-resolution scans obtained later in the mission will eventually test this hypothesis and also reveal how the feature evolves.

It can be affirmed that the observations presented above constitute new and persuasive evidence for the presence of strong atmospheric subsidence near the poles. This fits very well with the suggestion by Murray *et al.*⁷ and Suomi and Limaye⁸ of vortex-type circulation at the poles in the UV imaging sequences of Venus by Mariner 10. Suomi and Limaye⁸ state that these

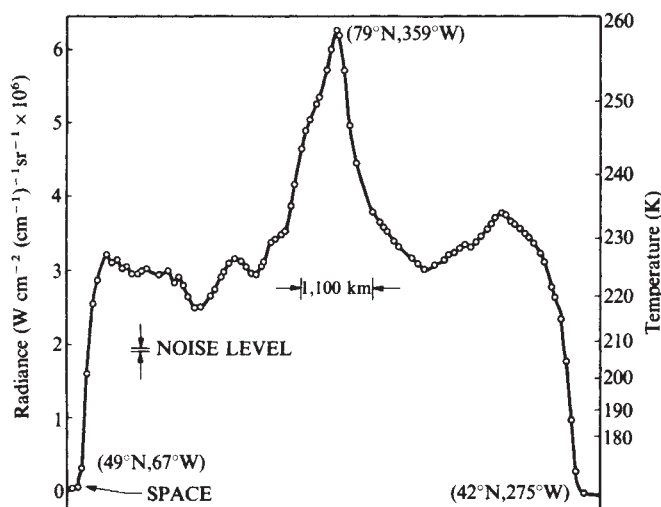


Fig. 1 A single scan across Venus at a wavelength of $11.5\ \mu\text{m}$ (spectral resolution = $0.027\ \mu\text{m}$) showing the effective cloud-top temperature observed. The scan cuts through the cold polar collar, showing its asymmetry, and through the anomalously hot feature near to, but offset from, the north pole. The scan corresponds to the bottom edge of the image in Fig. 3.

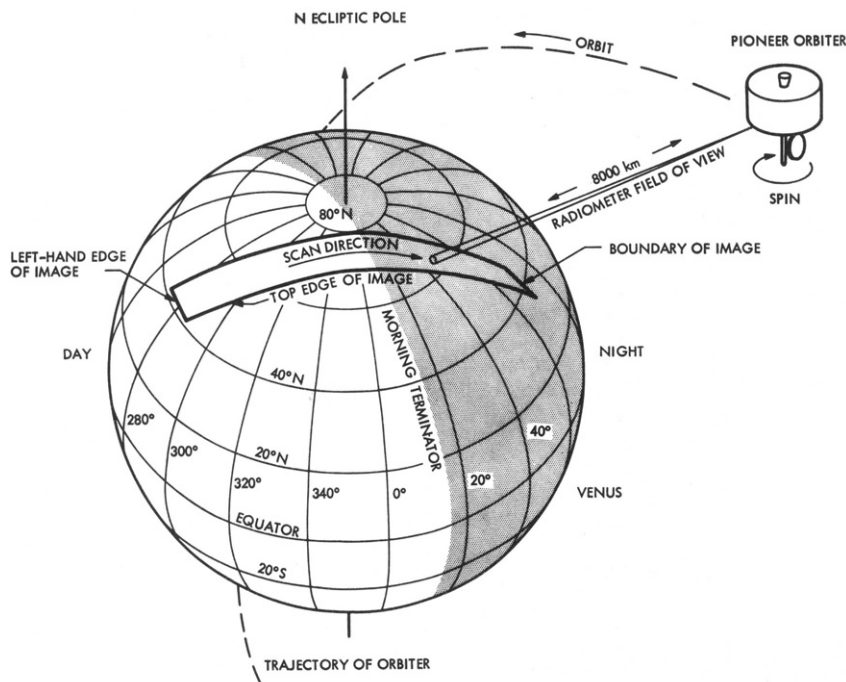


Fig. 2 Diagram showing the geometry of the observations, and the boundaries of the high-resolution image.

observations of spiral bands of cloud and a poleward component of cloud motion "are compelling suggestions that at least during the 7 days of the Mariner 10 flyby in 1974 the stratospheric circulation was composed of two giant vortices more or less centred on each pole with meridional inflow from low latitudes towards each pole. The vortex... would be characterised by a region of mass sink in the polar regions in the upper atmosphere and a mass source at lower latitudes, essentially a hemispheric Hadley circulation cell strongly organised by the vertical zonal flow".

The 'single cell' model for the Venus stratosphere is attractive, not least because it provides an intuitive explanation for the ubiquitous and uniform cloud cover outside the polar domains. It is easy to imagine that the atmosphere may be slowly rising over most of the area of the planet, thus sustaining dynamically the observed cloud deck. This mass flow would then be balanced by relatively rapid descent within the polar vortices. There are difficulties with this, however. First, it is difficult to explain why the 'eye' of the vortex is not centred on the rotational pole. The low resolution data alluded to earlier show clearly that its centre is 5°–10° from the pole itself. Second, the presence of the thick, high polar collar cloud (segments of which can be seen in Fig. 3) is not explained by a simple vortex circulation and may, in fact, preclude it. Third, the downwelling in the polar region must be explained in terms of a circulation mechanism which produces a poleward meridional flow⁸, but against the gradient of increasing temperature from equator to pole⁴. Such a situation is not

consistent with a classical 'Hadley' type circulation driven by solar heating. The stratospheric dynamics on Venus require some more complex mechanism, perhaps involving motions driven from below.

This is one phase of work carried out at the Jet Propulsion Laboratory, California Institute of Technology, under contract NAS 7-100 sponsored by NASA. D.J.D. is supported by a NASA/NRC Resident Research Associateship.

FREDRIC W. TAYLOR
DANIEL J. MCCLEESE
DAVID J. DINER

*Earth and Space Sciences Division,
Jet Propulsion Laboratory,
California Institute of Technology,
Pasadena, California 91103*

Received 3 April; accepted 9 May 1979.

1. Murray, B. C., Wildey, R. L. & Westphal, J. A. *J. geophys. Res.* **68**, 4813–4818 (1963).
2. Diner, D. J., Westphal, J. A. & Schilver, F. P. *Icarus* **27**, 191–196 (1976).
3. Apt, J. & Goody, R. *Science* **203**, 785–787 (1979).
4. Taylor, F. W. *et al. Science* **203**, 779–781 (1979).
5. Taylor, F. W. *J. Atmos. Sci.* **32**, 1101–1105 (1975).
6. Ksanfomaliti, L. V. *et al. Sov. Astr.-AJ.* **20**, 476–481 (1976).
7. Murray, B. C. *et al. Science* **183**, 1307–1315 (1974).
8. Suomi, V. E. & Limaye, S. S. *Science* **201**, 1009–1011 (1978).



Fig. 3 An 11.5- μ m image of thermal emission from the region near the north pole of Venus delineated in Fig. 2. The data were obtained from the Pioneer orbiter on 15 December 1978. The brightest area is emitting more intensely, and hence is warmer, than its surroundings. The dark (cold) circumpolar collar appears either side of the hot region, and some of its detailed structure can be seen. The limb of the planet is visible at either edge of the frame.

Evidence for lightning on Venus

WHETHER lightning exists in a planetary atmosphere is a fundamental question for the atmospheric physicist. The conditions within a lightning stroke are far different from those within ambient atmosphere, permitting chemical and physical changes in the atmosphere which are not possible in equilibrium conditions. The relative importance of such nonequilibrium processes depends on the frequency and the location of this electrical activity. Discovery of lightning on other planets would also affect other scientific fields, including studies of planetary magnetospheres and of life. Lightning has been predicted or considered on Venus¹ and Jupiter^{2,3}, and evidence for lightning has been obtained with instrumentation on Venera 11 and 12 descent vehicles⁴. Electric fields characteristic of lightning, and acoustic signals, were detected on the dayside of the planet. We

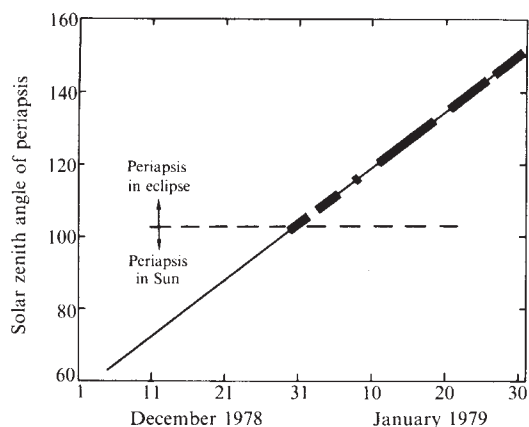


Fig. 1 The solar zenith angle of the periapsis of the Pioneer Venus Orbiter in December 1978 and January 1979 with orbits with impulsive events marked. The events, possibly due to lightning in the atmosphere of Venus, seem to be observed only when the Orbiter is in the shadow of Venus.

present here evidence of lightning on Venus obtained by instruments on the Pioneer Venus Orbiter.

The spacecraft was placed into Venus orbit on 4 December 1978⁵. Results discussed here are derived from observations to 31 January 1979. The periapsis of the Orbiter is well within the Venus ionosphere where electromagnetic waves characteristic of lightning might propagate. Initially periapsis was in sunlight. However, periapsis moved into darkness by the end of December, and at this time the Orbiter electric field detector obtained its first evidence for lightning.

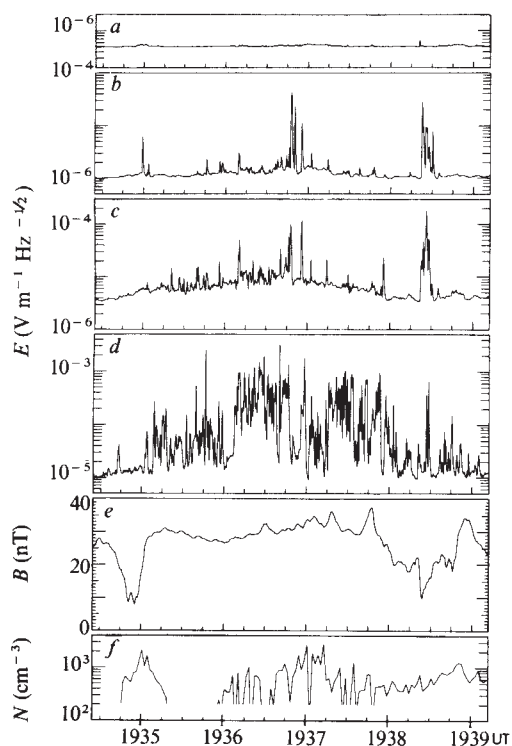


Fig. 2 A few minutes of electric and magnetic field data observed with Pioneer Venus Orbiter on 21 January, 1979. The measured electric field amplitudes of waves with frequencies *a*, 30 Hz; *b*, 5.4 kHz; *c*, 730 Hz; *d*, 100 Hz. The magnitude of the magnetic field is shown in *e*. Electron density data are shown in *f* (unconnected for densities $< 500 \text{ cm}^{-3}$). Periapsis was at 19 h 36 min 32 s.

The electric field detector has sensors which are mounted on the body of the Orbiter spacecraft with an effective antenna length of 0.7 m. The amplitude of the voltage induced on the antenna is measured in four frequency bands centred at 0.100, 0.730, 5.40 and 30.0 kHz, each 30% wide. The amplitudes are telemetered at various rates, commonly a complete spectrum (four measurements) every 0.5 s. The data reported here are from three instruments on Pioneer Venus—the electric field detector⁶, the magnetometer⁷, and the Langmuir probe⁸.

Electric field data from early orbits were carefully inspected for evidence of lightning but none was found. In fact, in the dayside ionosphere, the wave levels were particularly low during these early orbits, especially in the two lower bands. Beginning with periapsis data on 30 December 1978, strong, very impulsive signals were often observed at low altitudes. Figure 1 shows the precession of the solar zenith angle of the periapsis of the orbit with the days of the low altitude impulsive events indicated. Reception of the waves at the satellite appears possible only when the satellite is in or near the shadow of Venus. In principle, this could be due to a source effect (for example, if venusian thunderstorms occur only at night) or to a propagation effect (for example, if the daytime venusian ionosphere is opaque to the waves). A source effect can probably be ruled out since Venera 11 and 12 lightning observations were made at solar angles $< 25^\circ$, and this fact alone suggests that the nightside observations on the Orbiter are associated with changes in the ionosphere.

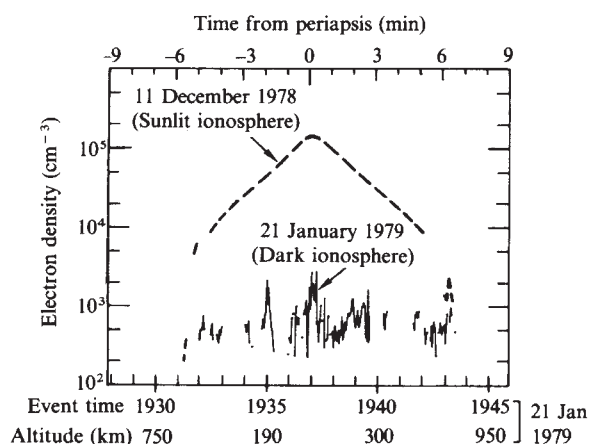


Fig. 3 Onboard measurements of the electron density for a typical dayside and nightside ionosphere (dashed line) and for the nightside ionosphere corresponding to the data shown in Fig. 2 (solid line).

Figure 2 shows a fairly typical example from 21 January 1979 when the altitude of periapsis was 164 km (at 19 h 36 min 32 s). The impulsive wave events were primarily detected at altitudes less than about 250 km, with regions of low electric field activity in the ionosphere above about 250 km. The magnetic field was about 30 nT during most of the time shown in Fig. 2, giving an electron gyrofrequency of about 840 Hz. Within the density 'dropouts' we use $N \approx 40 \text{ cm}^{-3}$ as an electron density estimate. For a magnetic field of 30 nT and an electron density of 40 cm^{-3} , 100 and 730 Hz waves will propagate in the whistler mode, whereas no cold plasma wave mode will support propagation at 5.4 or 30 kHz. The events are strongest at 100 and 730 Hz although a few events extend to 5.4 kHz as well. The 5.4 kHz data may be due to propagation in other wave modes or may be due to waves leaking out of a nearby region of depleted electron density (common on the nightside of Venus). Figure 3 shows the different character of the ionosphere on the day and night sides of Venus. The dayside is relatively more dense and more regular. The impulsive events occurred at an average rate of about 0.5 s^{-1} between 19 h 35 min and 19 h 39 min. The actual

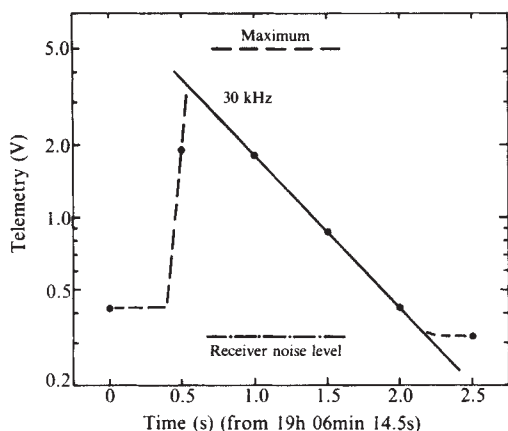


Fig. 4 The decay of the measured amplitude of an impulsive event on 13 January, 1979. The decay is exponential and apparently due entirely to the electronic averaging in the instrument. This indicates that the length of the input signal was very much less than the time constant of 0.70 s.

rate may have been much higher since the time between measurements was 0.5 s during this period. Venera 11 observed a maximum impulse rate of 25 s^{-1} (ref. 1).

Lightning on Earth is very impulsive. A test of the impulsiveness of the events is to observe the decay of the receiver output after an isolated impulse. The pulse shape of such an isolated impulse on 13 January 1979 is shown in Fig. 4. The rise time of the pulse is very short but, as expected, the decay is exponential with a time constant of 0.70 s, consistent with the decay time constant of the receiver.

The following evidence leads to the tentative conclusion that the impulsive events were caused by venusian lightning: (1) The signals are intense and highly impulsive; (2) the signals are detected near periapsis, well inside the ionosphere; (3) the spectral characteristics of the signals are generally consistent with whistler wave propagation up through the ionosphere; (4) the signals are often observed during intervals when low and variable electron densities are measured.

When the nightside densities are as high as 10^3 cm^{-3} , the interpretation that the impulsive events were caused by lightning requires that the whistler mode damping that seems to be effective in the dayside ionosphere⁶ should be ineffective in the lower density nightside ionosphere. Other interpretations of some of the wave measurements are also conceivable (for example, some observations could represent electrostatic ion acoustic or electrostatic ion cyclotron waves^{9,10}, or be due to turbulence near zero frequency¹¹). However, many of the impulsive events are detected near regions with local electron densities $< 10^2 \text{ cm}^{-3}$, and since the electron energy for Landau damping is proportional to the inverse of the density, the cool electrons near periapsis at night can never provide much damping. Moreover, for densities $< 10^2 \text{ cm}^{-3}$ and a magnetic field of 30 nT, a 100-Hz whistler mode wave has a wavelength of 10 km or more, and it is almost certain that the spacecraft is frequently as close as the wavelength from the lower edge of the ionosphere. For these cases, and others where ducts might be present, signals from atmospheric lightning should be able to propagate to the spacecraft without any significant damping.

This work was supported by NASA.

WILLIAM W. L. TAYLOR
FREDERICK L. SCARF

Space Sciences Department,
TRW Defense and Space Systems Group,
One Space Park,
Redondo Beach, California 90278

CHRISTOPHER T. RUSSELL

Institute of Geophysics and Planetary Physics,
University of California, Los Angeles,
405 Hilgard,
Los Angeles, California 90024

LARRY H. BRACE

NASA Goddard Space Flight Center,
Greenbelt, Maryland 20771

Received 21 February; accepted 19 April 1979.

- Otroshchenko, V. A. & Surkov, Yu. A. *Origins of Life* **5**, 487-490 (1974).
- Sagan, C. E., Lippincott, E. R., Dayhoff, M. O. & Eck, R. V. *Nature* **213**, 273-274 (1967).
- Bar-Nun, A. *Icarus* **24**, 86-94 (1975).
- Ksanfomaliti, L. V., Ganpantserova, D. F., Petrova, E. V. & Souvorov, A. P. *Z. Pisma Astr.* **5**, 5 (1979).
- Colin, L. *Science* **203**, 743-745 (1979).
- Scarf, F. L., Taylor, W. W. L. & Green, I. M. *Science* **203**, 748-750 (1979).
- Russell, C. T., Elphic, R. C. & Slavin, J. A. *Science* **203**, 745-748 (1979).
- Brace, L. H. *et al. Science* **203**, 763-765 (1979).
- Rodriguez, P. & Gurnett, D. A. *J. geophys. Res.* **80**, 19-31 (1975).
- Mozer, F. S. *et al. Phys. Rev. Lett.* **38**, 292 (1977).
- Temerin, M. *J. geophys. Res.* **83**, 2609-2616 (1978).

Observation of magnetic flux ropes in the Venus ionosphere

MAGNETIC field measurements made by the Pioneer Venus Orbiter spacecraft reveal a very low average field strength within the dayside ionosphere of Venus, typically only a few nanoteslas, in contrast to fields of several tens of nanoteslas just outside the ionosphere¹. Thus, at least in the range of solar zenith angles (65-90°) initially probed by the orbiter, the compressed interplanetary magnetic field of the shocked solar wind plasma (the magnetosheath) is effectively shielded from the ionosphere by currents flowing on the ionopause, the boundary between the ionosphere and the magnetosheath. In addition, the magnetic field pressure just outside the ionopause approximately balances the ionospheric thermal plasma pressure^{2,3}. However, within this generally low-field region the spacecraft occasionally passes through regions of very large field strength which can sometimes exceed that observed external to the ionosphere. These intense, short-lived enhancements are described here and interpreted to be due to the passage of the spacecraft through 'flux ropes', bundles of twisted magnetic field lines surrounded by ionospheric plasma.

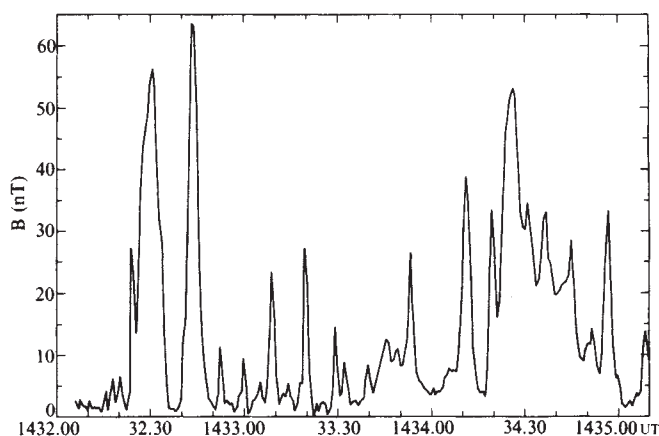


Fig. 1 Magnetic field enhancements observed within the ionosphere shortly after periapsis (at 1431.56) on orbit 3 of Pioneer Venus. Before and after this interval, the characteristic ionospheric field strength is less than 10 nT, while the peak field just outside the ionopause is ~55 nT.

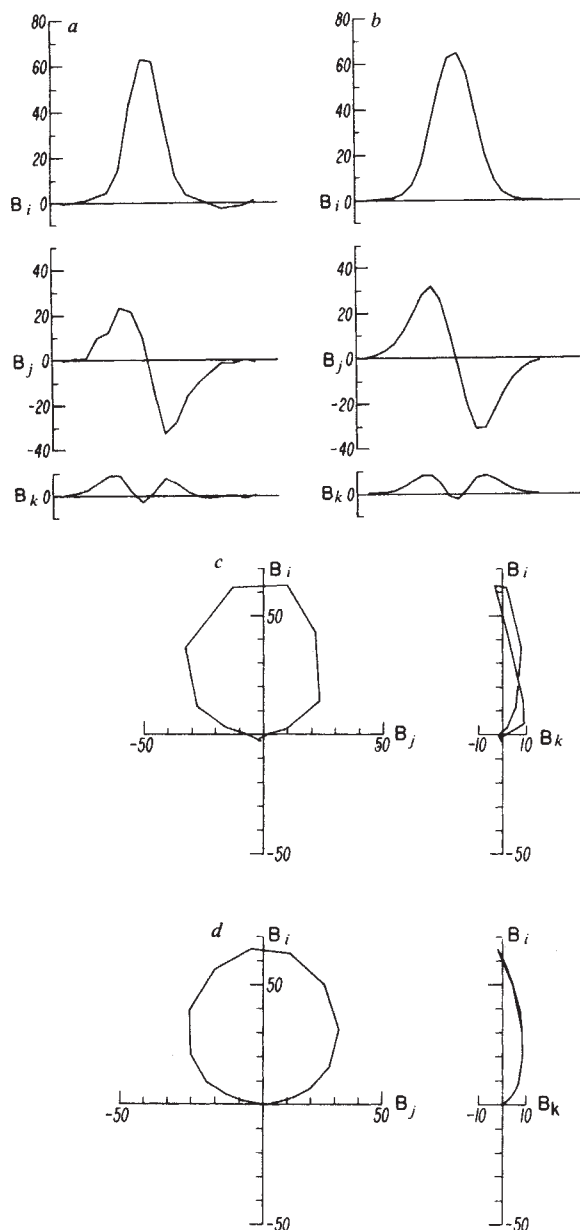


Fig. 2 *a, b* Comparison of the magnetic structure of one of the enhancements shown in Fig. 1 (*a*), and a model of the enhancement (*b*), the coordinate system is the principal axis system given by minimum variance analysis, and the fields are in nanoTeslas. The model parameters are discussed in the text. 13.5 s of data are shown. Hodograms of the structure (*c*) and the model (*d*). Orbit 3 of Pioneer Venus; 7 December 1978; 1432.37.3–1432.50.8; $l_A = 3.2$; $l_B = 4.5$; $\rho_{\min} = 2$.

Figure 1 shows the magnetic field strength observed near periapsis on the third orbit of Pioneer Venus. The spacecraft moves from an altitude of 209 km and a solar zenith angle of 66.7° to an altitude of 332 km and an angle of 69° during this time. Numerous enhancements occur throughout the ionospheric passage, and the field components (not shown) undergo complicated variations within each enhancement. It is unlikely that these variations are purely temporal, as the spacecraft is travelling faster than plausible wave disturbances. For example, with local ionospheric densities $>10^4 \text{ cm}^{-3}$, temperatures of roughly 2,000 K (refs 2, 3), and a background field of roughly 2 nT, the Alfvén speed is $\sim 0.5 \text{ km s}^{-1}$, while the sound speed is $\sim 1 \text{ km s}^{-1}$. The orbital speed of the spacecraft near periapsis is

$\sim 10 \text{ km s}^{-1}$, so we are probably observing spatial structures. Because the spacecraft grazes these structures at various distances from their centres, the magnitude and duration of the enhancements are variable, as can be seen in Fig. 1. This figure also shows that structures often overlap one another, and for clarity, we have analysed only distinctly separate enhancements.

In an attempt to reduce the field variation to two dimensions, we have performed minimum variance analyses⁴ on some of these enhancements. Figure 2*a* shows the field variation, in the principal axis system, of the structure observed between 1432.37.3 and 1432.50.8 UT on orbit 3. The *i*, *j* and *k* components refer to the maximum, intermediate and minimum variance directions respectively. Figure 2*b* shows the results of a minimum variance analysis applied to a passage through a model field describable by the following equations:

$$B_\phi = B(\rho) \times \cos(\alpha(\rho)).$$

$$B_z = B(\rho) \times \sin(\alpha(\rho))$$

where

$$B(\rho) = B_0 \times \exp(-\rho^2/l_B^2)$$

and

$$\alpha(\rho) = \pi/2 \times (1 - \exp(-\rho^2/l_A^2))$$

These equations define a helical magnetic field, whose magnitude and pitch α , depend on the radial distance, ρ , from the axis of the structure. B_ϕ and B_z are, respectively, the azimuthal and axial components of the field. The scale lengths l_A and l_B determine how rapidly the pitch and magnitude vary with distance from the axis. For the case shown in Fig. 2*a, b* $l_A = 3.2$, $l_B = 4.5$, and the distance of closest approach to the axis was 2.0. Hodograms of both the observations and the model are shown in Fig. 2*c, d*; the field vector swings through a full 180° in the *i-j* plane, passing through maximum magnitude near 90° . The distinctive variation in B_k is a result of passing through the helical structure off-axis.

The magnetic configuration suggested by the data and described explicitly by the model is shown in Fig. 3. Because of the helical nature of the magnetic field, the structure resembles a rope of twisted magnetic flux tubes: the magnetic field is axial and at its peak strength in the centre of the structure, becoming weaker and more azimuthal with increasing distance from the axis.

The currents which produce this helical magnetic structure are shown in Fig. 4*a, b*, both in the cylindrical coordinates of the structure and in coordinates parallel and transverse to the magnetic field at each point. Note that throughout the modelled passage the axial current is always dominant.

To determine the source and evolution of these flux ropes, we next examine their amplitude and occurrence frequency with altitude. Figure 4*c, d* shows these quantities combined for orbits 3, 4, 6 and 7. The altitude range is divided into 25-km bins, and the total number of flux rope structures per bin is shown in Fig. 4*c*. Because of its orbit, the spacecraft spends more time in the low-altitude bins than in the high, and this effect can be seen as an increase in total number of structures per bin with decreasing altitude. The altitude interval 150–175 km was sampled only during orbit 7, and does not have an observing time comparable to the higher altitude bins. To correct for the varying observing time with altitude we have determined the ratio of the total time spent within flux ropes to the total time spent within an altitude bin, as shown in Fig. 4*d*. There seems to be a slight increase in

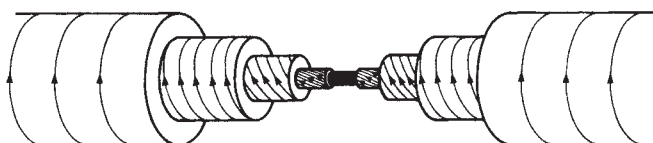


Fig. 3 Inferred magnetic structure of an ionospheric flux rope. The field is weak and azimuthal in the outer regions, becoming much stronger and more axial near the centre of the structure.

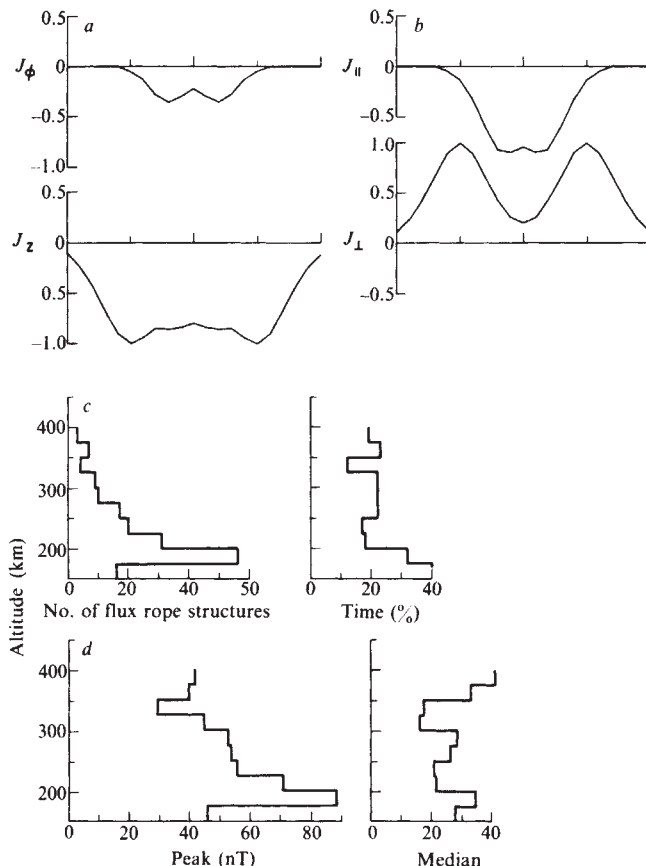


Fig. 4 *a, b* Current densities (arbitrary units) of the model structure in Fig. 2*a, b*. The cylindrical components (*a*) J_ϕ and J_z refer to the azimuthal and axial current contributions, while $J_\parallel$ and $J_\perp$ (*b*) are the contributions parallel and transverse, respectively, to the local magnetic field vector. The abscissa is distance along the spacecraft trajectory which in this case does not pass through the centre of the rope. Distributions of occurrence frequency (*c*) and amplitude (*d*) with altitude.

occurrence rate at lower altitudes. Additional data, especially when Pioneer Venus returns to the dayside ionosphere, will be needed to confirm this tendency.

If the magnetic field in these structures is in pressure balance with the surrounding ionospheric plasma, one would expect their characteristic field strength to increase with decreasing altitude, as the plasma thermal pressure increases. Figure 4*d* shows the peak flux rope amplitude observed within each altitude bin, and indeed, the amplitude seems to increase with decreasing altitude. The 150–175 km bin is the most notable exception to the trend, perhaps because this lowest altitude interval was sampled only on orbit 7. The right portion of Fig. 4*d* shows the median of the distribution of flux rope amplitudes within each bin. The distribution suggests that, at altitudes below 350 km, amplitudes increase overall with decreasing height, but results above 325 km have less statistical accuracy because of the small number of samples in those intervals.

Although the results shown are only based on four orbits of data they, nevertheless, carry implications for the source and evolution of flux ropes. First, the suggestion that these structures are more numerous at lower altitudes may be due to either a 'piling up' effect or to a source at low altitudes. The former might occur if magnetosheath magnetic flux tubes were pulled into the ionosphere by solar wind convection, such as has been proposed for the terrestrial magnetopause⁵. These structures would slow down and pile up in the denser, slower moving lower ionosphere. If, however, these structures result from a source in the

lower ionosphere, they might percolate upwards due to buoyancy; indeed, the enhanced magnetic field pressure within these structures implies the reduction of the plasma thermal pressure there, and if this shows up as a reduction in ion/electron number density, flux ropes must be buoyant with respect to the surrounding plasma. A similar process is believed to occur in the terrestrial equatorial ionosphere⁶. The helical nature of flux ropes could develop from velocity shear in the ionosphere, due perhaps to a viscous interaction with the magnetosheath at the ionopause. Given the observed ionospheric temperatures and densities, and flux rope scale diameters between 10 and 100 km, the magnetic diffusion time of such structures would be many hours. Thus flux ropes are likely to persist long enough to be transported, not only from the ionopause to the lowest ionospheric levels (or vice versa), but from the subsolar region to the terminator as well.

Whatever their source, these ionospheric flux ropes may play an important part in the solar wind–ionosphere interaction, and ionospheric plasma transport and heating. Analysis of the early Pioneer Venus data is continuing to define better the nature of this unexpected phenomenon.

C. T. RUSSELL
R. C. ELPIC

*Institute of Geophysics and Planetary Physics,
University of California,
Los Angeles, California 90024*

Received 26 February; accepted 10 April 1979.

1. Russell, C. T., Elphic, R. C. & Slavin, J. A. *Science* **203**, 745–748 (1979).
2. Brace, L. H. *et al.* *Science* **203**, 763–765 (1979).
3. Knudsen, W. C. *et al.* *Science* **203**, 757 (1979).
4. Sonnerup, B. U. O. & Cahill, L. J. *J. geophys. Res.* **72**, 171 (1967).
5. Lemaire, J. *Planet. Space Sci.* **25**, 887 (1977).
6. Woodman, R. F. & LaHoz, C. J. *geophys. Res.* **81**, 5447 (1976).
7. McClure, J. P., Hanson, W. B. & Hoffman, J. H. *J. geophys. Res.* **82**, 2650 (1977).

Polar heating and the shape of Venus

ON Venus, as on the other terrestrial planets, temperature variations are expected between equator and poles and between day and night. Several attempts have been made to measure these deviations from circular symmetry in the radio emission from Venus using interferometric techniques, for example, Berge and Greisen¹ at a wavelength of 3.1 cm and Sinclair *et al.*^{2–4} at 11 cm. The observed variations are found to be very small. Berge and Greisen¹ report a slight increase in the brightness temperature towards the poles, but the effect was less than one standard deviation of the measurements. A larger increase in brightness temperature towards the poles was seen by Sinclair *et al.*⁴. We report here a similar polar brightening at a wavelength of 6 cm which persuades us that the effect is real. This brightening is a surface or near surface phenomenon as at 6 cm about half the flux is from the surface and half from the lower atmosphere.

Our observations were made with a two-element interferometer at Owens Valley Radio Observatory between 6 May and 13 May 1977, near inferior conjunction with Venus. Each element of the interferometer is a steerable, parabolic, 27-m diameter antenna with a linearly polarised feed that can be rotated by remote control. The polarisation of the E-vector accepted from Venus was either perpendicular or parallel to the baseline projected on the planet and was alternated between successive records. The observational records were 5 min long with calibration records every 45–60 min. Small-diameter extragalactic radio sources with accurately known positions were used to calibrate the fringe amplitude and phase.

An east–west (E–W) baseline of 396 m and a north–south (N–S) baseline of 456 m were selected to observe near the first

null of the planet's 6-cm visibility function, which is the Fourier transform of the spatial brightness temperature distribution. These baselines were nearly parallel to N-S and E-W directions on Venus, as well as on the Earth. The resulting visibility functions are most conveniently expressed as a function of β , the instantaneous projected baseline length in wavelength times the angular radius of the planet based on a physical radius of 6,100 km. The value of β at the first null of the visibility function, β_0 , varies inversely with the width of the brightness-temperature distribution along the baseline. Thus, by comparing the E-W and N-S values of β_0 , the directions in which the planet is wider and/or hotter towards the limbs can be determined. We cannot uniquely deconvolve variations in the planet's radius and temperature variations. Our observations were all near the first null and determine β_0 quite accurately.

An interferometer measures both the amplitude of the visibility function and the phase. Theoretically, phase data can be used to detect asymmetries about a line through the centre of the disk perpendicular to the baseline. However, our phase data from the N-S baselines were affected by antenna pointing errors caused by strong winds, and we could not determine the amount of asymmetry between the north and south poles on Venus. To a close approximation, such pointing errors do not affect the measurements of β_0 . Sinclair *et al.*⁴ report that the pole-to-pole asymmetry is <3 K at a wavelength of 11 cm. Consequently, we have assumed that equator-pole variation is symmetric, that is, that the brightening occurs equally at both poles. Observations on our E-W baseline were not affected by pointing errors as the winds at this time were low.

Figure 1 shows our measurements with polarisation perpendicular to the baseline for the E-W and N-S baselines. To find the best fit value for β_0 , we used least squares to fit the quadratic formula, $B(\beta - \beta_0) + C(\beta - \beta_0)^2$, to the data for a series of values of β_0 . Because a quadratic is a good approximation of the true visibility function over only a limited range of β , we consider only data points with $0.56 < \beta < 0.68$ for the perpendicular polarisation. We determine the value of β_0 that minimises the sum of the squares of residuals:

$$S = \sum_i (f_i - [B(\beta_i - \beta_0) + C(\beta_i - \beta_0)^2])^2$$

where f_i = data point at β_i and B and C are the best fitting coefficients for each value of β_0 . The best fitting values for β_0 are 0.6207 ± 0.0005 for N-S₁ and 0.6214 ± 0.0004 for E-W₁, where the errors are standard deviations determined from the

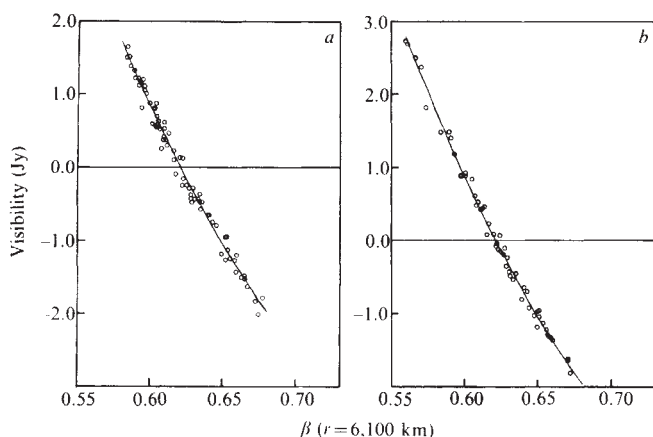


Fig. 1 Visibility data for *a* N-S and *b* E-W baselines with polarisation perpendicular to the baseline. *a*, $\beta_0 = 0.6207 \pm 0.0005$; *b*, $\beta_0 = 0.6214 \pm 0.0004$. The visibility is plotted plus or minus according to the phase. The curves shown and the values given for β_0 were determined by a least-squares fit of a quadratic as described in the text. The radius used in calculating β was 6,100 km.

data. The difference in nulls $\Delta\beta_0$, is 0.0007 ± 0.0006 , where the error again is the standard deviation.

Our measurements with polarisation parallel to the baseline are shown in Fig. 2. The best fit values over the range $0.55 < \beta < 0.67$ for β_0 are 0.6075 ± 0.0005 for N-S₁ and 0.6082 ± 0.0004 for E-W₁. The parallel nulls are at smaller values of β_0 than the perpendicular nulls because of a maximum in the brightness in this polarisation at the Brewster angle. The difference in nulls for the parallel polarisation is also 0.0007 ± 0.0006 .

As the measurements in the two polarisations are statistically independent, we may combine our results to get $\Delta\beta_0 = 0.0007 \pm 0.0004$. Note that in both polarisations the N-S β_0 is less than the E-W β_0 . This difference implies that the planet is either larger or brighter at the limb along the N-S direction than along the E-W direction. This result is surprising. If we assume that the $\Delta\beta_0$ is produced by a variation in radius of a uniformly bright planet, the planet's figure must be $7 \text{ km} \pm 4 \text{ km}$ prolate. If the planet's gravitational equipotentials are prolate to this extent as well, the resulting J_2 is about -10^{-3} , but J_2 for Venus has been estimated to be $-2 \pm 5 \times 10^{-6}$ by Esposito (personal communication) and $4.0 \pm 1.5 \times 10^{-6}$ by Akim *et al.*⁵.

It is difficult to compare our results with the polar heating reported by Sinclair *et al.*⁴, because they do not report a model-independent measure of the effect such as the prolateness of an equivalent, uniformly bright elliptical disk. Sinclair *et al.* altered their standard circularly symmetric Venus model to include a latitudinal surface temperature variation of the form $\Delta T \sin^2 \varphi$ where φ is the latitude. The form $\sin^2 \varphi$ was chosen as it is the first even term of a harmonic expansion. The resulting surface temperature at each latitude was adiabatically extrapolated upwards through the atmosphere. Visibility functions computed from this model were compared with their measured visibilities and $\Delta T = 16 \pm 2.5 \text{ K}$ was estimated. Apparently, the quoted error does not include any model uncertainties. A similar procedure with our data would yield a ΔT of about 5 K. The magnitude of the effect may be variable.

There are several viable phenomena that would cause the poles of Venus to be hot or to appear hot at radio wavelengths, although the thermal time constant in the lower venusian atmosphere is very long and, on theoretical grounds, one expects nearly constant temperatures on any equipotential surface. The absence of appreciable E-W asymmetries in the radio brightness across the terminator⁴ strongly supports this idea. However, the global atmospheric circulation may cause relative heating towards the poles. There may be strong downwelling causing warming associated with the circumpolar vortex⁶. IR measurements on the Pioneer Venus Spaceprobe⁷ detected a brightening of $\sim 10 \text{ K}$ over the north pole, down to an altitude level of about 70 km. These measurements are only sensitive to the atmospheric temperatures above the clouds. This region contributes nearly zero flux at radio wavelengths. It is unlikely that these elevated temperatures above 70 km are related to the brightening at the surface. A second possible cause of polar heating is a slight topographical flattening relative to the gravitational equipotentials possibly due to a more rapid spin state in the past or to the dynamic alignment of the planet's rotation axis and its axis of greatest moment of inertia. The latter case would apparently also require a more rapid rotation in the past. It is also possible that topography generated by geological processes could coincidentally cause this effect. Two phenomena that would increase the radio brightness towards the poles without increasing the physical temperature are a systematic decrease in radio opacity caused by a thinning of the clouds towards the poles or a systematic increase in surface emissivity away from the equator due to latitudinal differentiation of the surface geology.

The possibility of a greater topographic than gravitational oblateness is explored below. Atmospheric circulation will be considered elsewhere. It is not unreasonable to expect Venus's figure to have a small excess topographic oblateness. The Earth's figure has an excess oblateness of about 0.5 km, primarily because a majority of the continents are located nearer

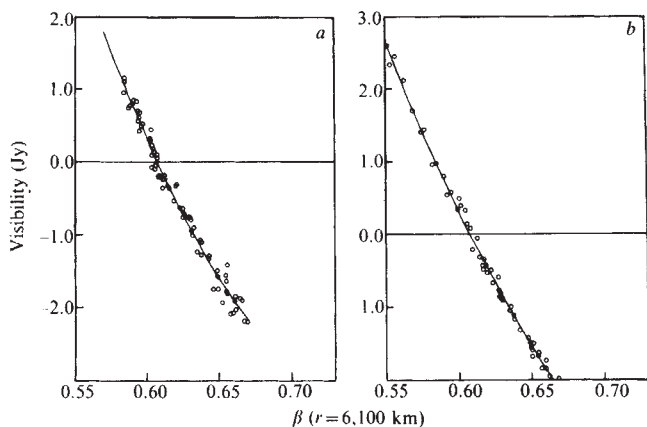


Fig. 2 Visibility data for *a* N-S and *b* E-W baselines with polarisation parallel to the baseline. *a*, $\beta_0 = 0.6075 \pm 0.0005$; *b*, 0.6082 ± 0.0004 . Otherwise as Fig. 1.

the Equator than the poles⁸. The Moon and Mars also have excess oblateness of about the same magnitude^{9,10}.

Model calculations were carried out to determine how large a bulge is required to explain our observations. In the model, the figure of the planet is assumed to be an oblate spheroid. In general, the gravitational equipotentials are oblate also due to the oblate figure and rotation. However, if the planet is not in hydrostatic equilibrium, the topographical oblateness will exceed the gravitational oblateness unless the density distribution is prolate. For simplicity in calculations, the gravitational equipotentials in the model are assumed to be spheres. The atmosphere is well mixed so the surfaces of constant temperature and pressure are also spherical. The poles are lower in the atmosphere than the equator and are, therefore, warmer. An atmospheric model and a computational method similar to that of Muhleman *et al.*¹¹ are used to calculate brightness temperatures which are integrated numerically over the planet's disk to find the model visibility function for an E-W and a N-S baseline. In calculating brightness temperatures, the effects of absorption, emission and refraction in the atmosphere are considered along with the effects of reflection and emission by a dielectric surface. We find that a topographical oblateness of $0.6 \text{ km} \pm 0.4 \text{ km}$ will yield a $\Delta\beta_0$ that agrees with our observations. We believe that this is a good approximation of the amount by which the topographical oblateness must exceed the gravitational oblateness. This amount of oblateness together with the $7.8^\circ \text{ km}^{-1}$ lapse rate leads to the poles being $5^\circ \pm 3^\circ$ warmer than the equator.

To determine whether the excess topographical oblateness could be due to a fossil rotational bulge, we use the model of Melosh¹² for a despun planet. The structure of the planet in this model is an elastic shell over a fluid interior. We assumed an initial rotation period of 1 day and a density contrast between the interior and the elastic shell of $1.5\text{--}2.0$. We find that when the excess topographical oblateness is $0.6 \pm 0.4 \text{ km}$, the gravitational field generated has a J_2 an order of magnitude above the observational limit. These results indicate that the presence of a fossil rotational bulge on Venus large enough to explain our observations is unlikely. However, the polar heating could still be due to regional isostatically compensated changes in the topography which would have less impact on J_2 .

In conclusion, our observations are consistent with previous work that has shown that the disk of Venus is nearly circularly symmetric at radio wavelengths. However, a small amount of polar brightening is present. We cannot yet determine the cause of this phenomenon although we have argued against the possibility of a fossil rotational bulge large enough to explain the result. As the poles of Venus have other interesting aspects involving atmospheric temperatures and circulation, they are an area worthy of future investigation.

This work was supported by NASA grants NGR 05-002-114 and NGL 05-002-003 and NSF grant AST 00247. J.B.P. is partly supported by an NSF graduate research fellowship.

J. B. PECHMANN*

*Division of Geological and Planetary Sciences,
California Institute of Technology,
Pasadena, California 91125*

D. O. MUHLEMAN
G. L. BERGE

*Owens Valley Radio Observatory,
California Institute of Technology,
Pasadena, California 91125*

Received 6 February; accepted 20 April 1979.

*To whom correspondence should be addressed.

1. Berge, G. L. & Greisen, E. W. *Astrophys. J.* **156**, 1125-1134 (1969).
2. Sinclair, A. C. E., Basart, J. P., Buhl, D., Gale, W. A. & Liwshitz, M. *Radio Sci.* **5**, 347-354 (1970).
3. Gale, W. A. & Sinclair, A. C. E. *Astrophys. J.* **175**, 535-553 (1972).
4. Sinclair, A. C. E., Basart, J. P., Buhl, D. & Gale, W. A. *Astrophys. J.* **175**, 555-572 (1972).
5. Akim, E. L., Vlasova, Z. P. & Chulko I. V. *Sov. Phys. Dokl.* **23**, 313-315 (1978).
6. Suomi, V. E. & Limaye, S. S. *Science* **201**, 1009-1011 (1978).
7. Taylor, S. W. *et al. Science* **203**, 779-781 (1979).
8. Balmino, G., Lambeck, K. & Kaula, W. M. *J. geophys. Res.* **78**, 478-481 (1973).
9. Bills, B. G. & Farrari, A. J. *Icarus* **31**, 244-259 (1977).
10. Bills, B. G. & Farrari, A. J. *J. geophys. Res.* **83**, 3497-3508 (1978).
11. Muhleman, D., Orton, G. & Berge, G. *Astrophys. J.* (submitted) (1979).
12. Melosh, H. J. *Icarus* **31**, 221-243 (1977).

X-ray, optical and radio observations of A1710-34

SAS 3 ROTATING modulation collimator (RMC) measurements are reported here which yield a new, more precise, position for the Ariel 5 X-ray source A1710-34 (ref. 1). We also report optical observations of stars in the X-ray error box and discuss evidence obtained with the Parkes 64 m radio telescope for 2-cm radio emission from the vicinity of the X-ray source.

Measurements were made of A1710-34 with the 2.3' and 4.5' FWHM RMC detectors on the SAS 3 X-ray observatory^{2,3}. These detectors are sensitive in the range 2-11 keV. Two independent sets of position measurements were made between 1600 UT 16 July and 1200 UT 22 July 1978. Details of the observations and the most probable position for the source (which we designate 2S1711-339) are given in Table 1. This position is 2.2 arc min from the most probable position determined using the Ariel 5 satellite and is therefore just inside the Ariel 5 error box. The 90% error circle of the new position has a radius of 40 arc s. A finding chart from the SRC Schmidt Survey J-plate is presented in Fig. 1.

The error circle for this observation of 2S1711-339 includes seven stars of the European Southern Observatory (ESO) quick blue survey plate with magnitude, $m_{\text{pg}} \leq 20$. The brightest star (number 1, Fig. 1) has a magnitude, $m_{\text{pg}} \sim 15$. The Anglo-Australian Telescope and the RGO spectrograph (140 Å mm⁻¹ dispersion) were used to obtain a spectrum (3,200-7,000 Å) of this star. The star is of spectral type G0 and has no perceived emission lines. We therefore have no spectroscopic evidence to suggest that it is the optical counterpart to the X-ray source.

The SRC-J survey plate shows that the X-ray error circle lies in the direction of a dust lane. Absorption is particularly strong in the eastern half of the error circle. If the X-ray low energy spectrum is heavily cut off, the optical candidate is more likely to be faint and to the east behind the dust lane. Low-energy X-ray observations would be useful to clarify this particular point. No particularly blue or red stars or variables could be identified by comparing ESO (blue) and SRC (yellow) survey plates.

Table 1 RMC observations of A1710–34 (2S1711–339)

Date (1978)	Exposure	Collimator	RA (1950)	Dec (1950)	Intensity μJy (2–11 keV)
0700 UT 20 July	49,500 s	2.3' FWHM	257.7579	–33.9897	13.2 ± 2.2 (6σ)
1200 UT 22 July			17 h 11 min 1.9 s	–33°59'23"	
0700 UT 20 July	49,500 s	4.5' FWHM	257.7500	–33.9978	14.6 ± 2.3 (7σ)
1200 UT 22 July			17 h 11 min 0.0 s	–33°59'52"	
1600 UT 16 July	74,000 s	2.3' FWHM	257.7583	–33.9908	16.6 ± 2.6 (7σ)
0500 UT 20 July			17 h 11 min 2.0 s	–33°59'27"	
1600 UT 16 July	74,000 s	4.5' FWHM	257.7554	–33.9908	15.3 ± 2.2 (7σ)
0500 UT 20 July			17 h 11 min 1.3 s	–33°59'27"	
Average			257.7554	–33.9922	14.9 ± 1.2
			17 h 11 min 1.3 s	–33°59'32"	

Table 2 Summary of X-ray, radio and optical observations of A1710–34

Waveband	Position (1950.0)		Error circle radius (90% confidence)	Intensity
	RA	Dec		
X-ray (2–11 keV)	17 h 11 min 1.3 s	–33°59'32"	40 arc s	$14.9 \pm 1.2 \mu\text{Jy}$
Radio (2 cm)	17 10 52	–34 00 36	2.2 arc min	$23 \pm 5 \text{ mJy}$ (max)
Optical (star 1)	17 11 1.32	–33 59 12.8	2 arc s	$m_{\text{pg}} \sim 15$

The region of the sky in the vicinity of A1710–34 (2S1711–339) has been searched for 2-cm radio emission using the Parkes 64m radio telescope with a full width to half-power beam of 2.2 arc min. Dual-beam switching between two identical beams separated by 6.5 arc min was used. The measurements were made in June 1977, August 1977 and April 1978 during a comprehensive search⁴ for new radio counterparts of X-ray sources using Ariel 5, SAS 3 RMC and HEAO 1 positions. The new X-ray position was not known at the time of these observations.

No significant signal was detected from the vicinity of A1710–34 (2S1711–339) in June 1977 or April 1978. Two-sigma upper limits were 66 and 12 mJy respectively. A flux density of $23 \pm 5 \text{ mJy}$ was detected in August 1977. As the

source was too weak for a reliable radio position determination, the beam was pointed at the Ariel 5 RMC source position. The sensitivity to a point source at the centre of the SAS 3 X-ray error circle would be $\sim 13\%$ of the on-axis response.

Radio counterparts of galactic stellar X-ray sources tend to be very variable. Thus, for example, the radio emission from Sco X-1 varies in an irregular manner with a typical time scale of hours. The range of variability is about two orders of magnitude⁵. GX9+1 and GX17+2 exhibit short term variability of the Sco X-1 type⁶. Circ X-1 is also variable on a time scale of hours⁷.

We have applied the χ^2 test to the hypothesis that the radio signal seen in August 1977 could have occurred by chance from a source of constant flux density equal to the weighted mean of the three observations. This gives $\chi^2 = 7.7$ for 2 d.f. The probability of obtaining a value of $\chi^2 \geq 7.7$ is about 2%. It therefore seems reasonable to assume that A1710–34 (2S1711–339) is a variable radio emitter with an intensity which rises above the threshold of the Parkes radio telescope at 2 cm only during periods of major flaring.

We conclude that there is probably a variable 2-cm radio source in the vicinity of A1710–34 (2S1711–339) but that further observations are required to determine the position accurately. In Table 2 we summarise the observations made in the three different wavelength bands.

We thank D. C. Morton for the spectroscopic observation of star 1. We acknowledge the support of the staffs of the Center for Space Research at MIT and the Goddard Space Flight Center. J.G.G. received financial support from Imperial College, London and MIT. Support has also been provided by the Australian Research Grants Committee.

J. G. GREENHILL
R. M. THOMAS
M. L. DULDIG

Physics Department,
University of Tasmania,
GPO Box 252C, Hobart,
Australia 7001

J. G. JERNIGAN
Department of Physics & Center for Space Research,
Massachusetts Institute of Technology,
Cambridge Massachusetts 02139

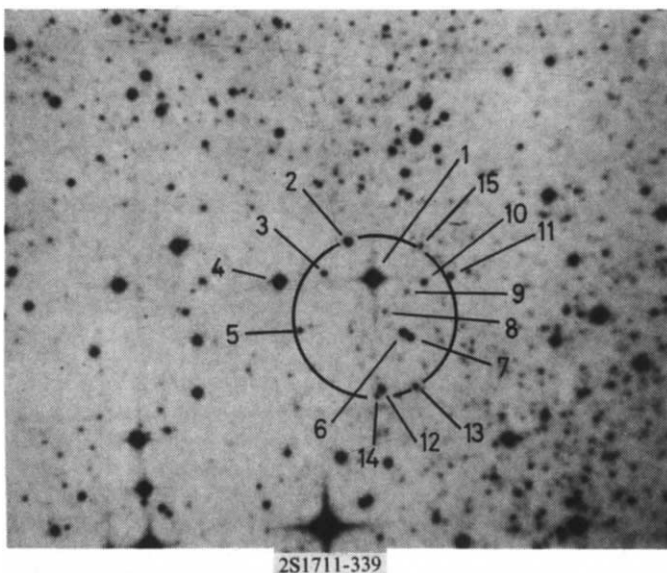


Fig. 1 Finding chart for the SAS 3 RMC position for 2S1711–339, taken from the SRC Schmidt Survey J-plate. The scale is given by the diameter of the error circle which is 80 arc in diameter. North-east is in the upper left-hand corner.

Anglo Australian Observatory,
PO Box 296, Epping,
Australia 2121

P. G. MURDIN

CSIRO Division of Radiophysics,
PO Box 76, Epping,
Australia 2121

R. F. HAYNES

Received 3 April; accepted 10 May 1979.

1. Carpenter, G. F., Eyles, C. J., Skinner, G. K., Wilson, A. M. & Willmore, A. P. *Mon. Not. R. astr. Soc.* **179**, 27P–34P (1977).
2. Doxsey, R. E., Apparao, K. M. V., Bradt, H. V., Dower, R. G. & Jernigan, J. G. *Nature* **269**, 112–116 (1977).
3. Bradt, H. V. *et al. Nature* **269**, 496–497 (1977).
4. Duldig, M. L. *et al. Mon. Not. R. astr. Soc.* **187**, 567–580 (1979).
5. Miyamoto, S. & Matsuoka, M. *Space. Sci. Rev.* **20**, 687–755 (1977).
6. Braes, L. L. E. & Miley, G. K. *IAU Symp. No. 55*, 86–97 (1973).
7. Whelan, J. A. J. *et al. Mon. Not. R. astr. Soc.* **181**, 259 (1977).

On the detection of jovian companions to white dwarfs

BRACEWELL¹ has recently proposed detection of planetary companions to nearby stars by means of a spinning IR interferometer. In particular, he notes that at far IR wavelengths, the emissivity ratio of the planet to the star increases by five orders of magnitude over equivalent values in the visible. I point out here that more than another three orders of magnitude in sensitivity is gained by choosing white dwarfs as the observational objects, which may permit detection of jovian or black dwarf companions by direct photometry from a space telescope with photon-limited, far IR detectors.

Compared with the Sun, a typical DA white dwarf has a surface area only $1\text{--}2 \times 10^{-4}$ as great, while its photospheric temperature is approximately double. Thus, at wavelengths in the Rayleigh–Jeans regime, an observer would detect 2,500–5,000 times fewer photons from a white dwarf than from a solar-type star at the same distance. The key assumptions to this proposal then are that planets of comparable or greater than jovian mass will (1) survive the parent star's red giant phase, and (2) presently contain internal energy sources similar to that of Jupiter. I discuss each of these assumptions in turn.

The most serious threat to a planet's survival is engulfment by the red giant photosphere. Viscous drag on the planet's orbital motion would then lead to a spiralling into the stellar core. The maximum luminosity² of a red giant with a $1.4 M_{\odot}$ carbon-oxygen core (that is the Chandrasekhar limit) is $\sim 5.2 \times 10^4 L_{\odot}$. If the red giant's effective temperature is $\geq 2,500$ K, all planets with orbital semi-major axes > 5.7 AU should remain outside the photosphere. For a $1.0 M_{\odot}$ core, the equivalent value is 4.2 AU. Runaway atmospheric evaporation might also cause planetary destruction. However, Jupiter's atmosphere must be heated to temperatures far above 3,000 K ($GM_p/kR_p \sim 180,000$ K) before significant quantities of atomic hydrogen will escape. Therefore, as long as a jovian planet remains outside the red giant photosphere, it will probably survive and perhaps even grow slightly from stellar wind accretion.

Both Jupiter and Saturn have internal energy sources. For Jupiter, this source provides more than half its observed thermal emission. Though gravitational contraction may be partially responsible, most astronomers favour an explanation relying on the radiation of primordial heat of formation. Graboske *et al.*³ have calculated the initial contraction and subsequent evolution of a $0.0095 M_{\odot}$ protoplanet and found that after a relatively short time ($\sim 10^6$ yr) spent along the Hayashi track, a rapid transition to the cooling curve of a very cold degenerate dwarf occurs. In this latter regime, $\log L/L_{\odot}$ scales as $-1.3 \log t$. As the average age of a white dwarf (including its progenitor's lifetime) is somewhat greater than half that of the Galaxy, the present

luminosity of a hypothetical jovian companion should remain at approximately half its value at $t = 4.6 \times 10^9$ yr, the age of our Solar System. Thus, one expects relatively strong internal emission to be a general property of all planets of jovian mass or greater.

The three or more orders of magnitude gained in relative sensitivity by the choice of white dwarfs rather than solar-type stars suggest that planetary companions might be detected as an 'IR excess' if advanced detectors become available. Adopting sample parameters of $\lambda = 40 \mu\text{m}$, $\Delta\lambda = 10 \mu\text{m}$, $R_{\text{WD}} = 0.013 R_{\odot}$, $R_p = 0.1 R_{\odot}$, $T_{\text{WD}} = 10^4$ K, $T_p = 120$ K, and a distance of 10 pc, a collecting area of 10^4 cm^2 would receive ~ 180 photons s^{-1} from the white dwarf and ~ 20 photons s^{-1} from the planet. A hypothetical photon-limited detector above the Earth's atmosphere could reveal the excess planetary signal with a signal-to-noise ratio > 10 in less than a minute. Confusion with the relic circumstellar dust shell of the red giant is unlikely because it should have been quickly swept away by the ram pressure associated with the white dwarf's motion through the interstellar medium. Zodiacal dust, however, may lead to serious background noise limitations¹ and increase detection times for a diffraction limited telescope by as much as a factor of 1,000 at low ecliptic latitudes.

One might also detect or set upper limits on the number of substellar, degenerate 'black dwarfs'⁴ in the mass range 10^{-3} to $\sim 0.06 M_{\odot}$ associated with white dwarfs. Such objects may provide an important contribution to the mass density, both locally and in clusters of galaxies^{5,6}. Their IR luminosity should scale at least linearly with mass. The advantage of a search in the immediate vicinity of a white dwarf is simply that one has a quite definite, highly localised position on the sky to measure. For a background-limited detection system, this allows one to maximise the signal-to-noise ratio until either the diffraction limit or the effective 'seeing' limit is reached.

There are over a dozen single white dwarfs known within 10 pc of the Sun⁷. A far IR survey of these stars from a space-borne telescope might quickly reveal whether solar system formation was a common event. A positive answer would give great impetus towards the construction of a more elaborate planet-detector, such as that proposed by Bracewell.

WILLIAM M. FAWLEY

Astronomy Department, University of Pennsylvania,
Philadelphia, Pennsylvania 19104

Received 12 February; accepted 23 April 1979.

1. Bracewell, R. N. *Nature* **274**, 780–781 (1978).
2. Paczynski, B. *Acta astr.* **21**, 417–435 (1970).
3. Graboske, H. C., Pollack, J. B., Grossman, A. S. & Olness, R. J. *Astrophys. J.* **199**, 265–281 (1975).
4. Kumar, S. S. *Astrophys. J.* **137**, 1121–1125 (1963).
5. Napier, W. McD. & Guthrie, B. N. G. *Mon. Not. R. astr. Soc.* **170**, 7–14 (1975).
6. Tarter, J. C. *Bull. Am. astr. Soc.* **8**, 517 (1976).
7. McCook, G. P. & Sion, E. M. *A Catalogue of Spectroscopically Identified White Dwarfs* (Villanova Univ. Obser., 1978).

Do comets provide material for the anomalous component of the cosmic rays?

MATTER from comets is proposed here to form a significant part of the source material for the 'anomalous component' of the cosmic rays. As a result it is concluded that certain molecular ions, such as CO^+ , should be present at energies around 10 MeV per nucleon.

The spectra of the nucleonic component of cosmic radiation in the energy band 1–60 MeV per nucleon have produced a number of surprises. The most dramatic of these concerned the fluxes of oxygen and carbon. Jokipii¹ gave minimum flux values in the region 5 MeV per nucleon for hydrogen and helium which essentially divided the solar particles on the low energy side from the true cosmic radiation at the higher energies. The

Chicago group reported in 1973² that the shape of the helium spectrum had changed in that the minimum reported earlier had now been filled in, whereas the hydrogen spectrum was essentially unchanged. Hovestadt *et al.*³ also detected in 1973 an anomalous flux of oxygen centred around 5 MeV per nucleon; as there was no corresponding flux of carbon the C/O ratio had changed by a considerable factor. Later experiments confirmed these measurements and showed the oxygen flux to be 20 times the value of the carbon flux at this peak. This is in complete contrast to the situation below 1 MeV per nucleon or above 60 MeV per nucleon where the C/O ratio remained close to unity. The later measurements also showed a substantial enhancement of nitrogen, and indicated a likely enhancement of neon. The total flux of the heavier elements is $\sim 1.0 \text{ m}^{-2} \text{ sr}^{-1} \text{ s}^{-1}$, in the band $3 < E < 60 \text{ MeV}$ per nucleon and that of helium is a few times larger.

The discovery of the anomalous component was very surprising and demanded an explanation. The generally accepted explanation is due to Fisk, Kozlovsky and Ramaty⁴ who postulate that these particles carry only a single charge in contrast to both the solar corpuscular radiation at $\sim 1 \text{ MeV}$ per nucleon and the cosmic rays at higher energies which are fully stripped. The consequently high magnetic rigidity ($\sim 2 \text{ GV}$) is an essential feature of their explanation. They envisage the following stages: (1) The passage of the Solar System through the local interstellar gas sweeps aside charged ions, while the neutral atoms are able to penetrate the magnetic cavity of the Sun. A fraction of these atoms is ionised well within the solar cavity. Species likely to be uncharged in interstellar space include He, N, O, Ne since their first ionisation potential is higher than that of the interstellar hydrogen which shields them. (2) After ionisation the singly charged species are carried by magnetic disturbances out to $\sim 50 \text{ AU}$ where magnetic inhomogeneities accelerate them preferentially, owing to their high rigidity. They retain their unit charge; the mean lifetime for further ionisation will be in the range 10^6 – 10^8 s , depending principally on their distance from the Sun. (3) After acceleration the ions, still presumably singly charged, diffuse through the Solar System, losing little energy from adiabatic deceleration. Again this is due to their rigidity being so much higher than cosmic ray and solar species of comparable energy that are fully stripped.

This model accounts for the particular enhancement of He, N, O and Ne atoms in terms of the magnitude of their first ionisation potential. The ions are introduced to the acceleration mechanism by being ionised well within the solar cavity. The onset of the anomalous component in 1973 and its continuation to the present is presumed to be concerned with the solar cycle and its effect on the acceleration mechanism.

If stages (2) and (3) operate on singly charged ions, clearly any other source of such ions might well contribute to the anomalous component. Comets seem to be such a source. The coma consists of neutral atoms and molecules which are the sublimation products from the cometary nucleus. The most abundant species in the coma seem to be degradation products of H_2O . In addition, in most comets, ionised atoms and molecules are swept by the solar wind into the familiar cometary tail. In this ion tail the spectrum from CO^+ is usually the most prominent component. Ions liberated by comets could be a direct source of material for stages (2) and (3) above. Further, atoms and molecules from the coma might be an alternative source of neutral particles for stages (1) to (3) above. The presence of molecules in this cometary debris leads us to an interesting possibility. These molecules might also be present in the anomalous component, as in many cases their lifetimes for further ionisation or dissociation have values comparable to those of atomic ions. A cometary source of material would have the following features.

First, the total amount of H_2O shed by short period comets is estimated by Delsemme⁵ to be $\sim 5 \times 10^7 \text{ g s}^{-1}$. In addition the amount of more refractory material released by comets to form meteoroids that produce the zodiacal light and the Gegenschein is estimated by Whipple and Huebner⁶ as 1 – $3 \times 10^7 \text{ g s}^{-1}$. These values are comparable with that coming in from the interstellar

Table 1 Binding energies of some abundant molecules and molecular ions

	Dissociation energy (eV)	Ionisation energy (eV)		Dissociation energy (eV)	Ionisation energy (eV)
H_2	4.54	15.43	H_2^+	2.71	16.3
CH	3.52	11.13	CH^+	3.65	17.2
C_2	6.27	12.0	C_2^+	5.53	16.8
CN	8.01	14.3	CN^+	4.96	19.5
CO	11.20	14.01	CO^+	8.45	22.1
NH	3.27	13.10	NH^+	3.77	18.3
N_2	9.88	15.58	N_2^+	8.83	23.4
NO	6.57	9.25	NO^+	10.93	25.6
OH	4.45	13.17	OH^+	4.88	18.5
H_2O	5.17	12.6	H_2O^+	5.75	19.3
CO_2	5.51	13.77	CO_2^+	5.35	19.0

medium as neutral atoms, which for oxygen is estimated as $\sim 5 \times 10^7 \text{ g s}^{-1}$, by Fisk *et al.*⁴.

Second, when first ionised the atom or molecule has a velocity and a component of momentum transverse to the magnetic field almost identical with the values possessed by the interstellar atoms.

Third, the position of the maximum release of cometary material is in the region of 1 AU, bracketed by the mean distances reached by neutral He, N and O before ionisation (0.5 – 4 AU)⁷.

Finally, the chemical composition of comets has several features in common with that of the anomalous particles. The detected spectra suggest that the principal components are degradation products of H_2O . However, contributions from the refractory material could include a greater variety of elements, for example, Mg and Si.

These points indicate that cometary material is apparently a good candidate for the source of at least some part of the anomalous component. The O/C ratio may place a constraint on the proportion of the anomalous component from comets, but cometary abundance values could well encompass the 20:1 ratio that one has from experiments on the anomalous component. One would not, however, expect a significant helium or neon contribution from a cometary source. The helium flux is well established, and has an abundance a few times that of oxygen, but the data for neon are dubious and are based on a very small number of detected particles, and could have been produced by incident molecular ions, such as CO, rather than neon. The presence therefore of inert gases in significant amounts implies that cometary material cannot be the only source of the anomalous component.

Finally, we would like to consider in further detail the possibility of molecular ions being present in the anomalous component. Table 1 lists some properties of appropriate molecules and molecular ions. Note that the stability of molecular ions against dissociation is comparable with their parent molecules, and their ionisation energy averages 1.5 times that of their parents. In practice the longevity of a molecular species in space is determined principally by whether any copiously produced excited states have a reasonably high dissociation probability. In O_2 , for example, UV absorption in the Schumann–Runge bands produces dissociation, which is a mechanism of considerable consequence, being the first stage of ozone production in the atmosphere. In contrast there are no such N_2 states abundantly produced and the effective threshold for N_2 is ionisation. The longevity of certain molecules, when essentially at rest at about 1 AU, has been estimated by Siscoe and Mukherjee⁸ and the figures computed are in the range 10^6 s .

As any molecular ion has one fewer weakly bound electron than its parent, and since its ionisation energy is ~ 1.5 times higher, its ionisation cross-section will be significantly smaller than that of its neutral parent. This in turn is smaller than the sum of the cross-sections of its atomic components. The cross-sections of atomic C, N, O, F and Ne are $\sim 10^{-17} \text{ cm}^2$, and that of

H is $\sim 3 \times 10^{-18} \text{ cm}^2$, for ionisation by protons and electrons having $\beta \sim 0.14$. Thus all the molecular ion species listed in Table 1 will have ionisation cross-sections close to 10^{-17} cm^2 when moving with $\beta \sim 0.14$. This leads to a partial ionisation rate at 1 AU of $\sim 4 \times 10^{-7} \text{ s}^{-1}$ (taking the density of solar wind as 10 cm^{-3}). At $\beta \sim 0.14$ the charge exchange cross-section with H^+ from the solar wind is very small, while the photoionisation cross-section for molecular ions will be smaller (because of the higher ionisation energies) than that of the parent molecules with $\beta \sim 0$ that have been considered by Siscoe and Mukherjee⁸. Thus the lifetime of the fast moving molecular ions against break-up is hardly dependent on its velocity and remains at $\sim 10^6 \text{ s}$ at 1 AU, increasing as the square of the distance from the Sun. We therefore conclude that the likely lifetime of suitable molecular ions, such as CO^+ , is adequate for a reasonable survival probability through the whole acceleration and diffusion stages proposed by Fisk *et al.*⁴. There is thus a reasonable chance of detectable fluxes of certain molecular ions with energies in the region of 10 MeV per nucleon.

From these considerations we feel that cometary material may have an important role as part of the source material for the anomalous component and that molecular ions from such a cometary source may be present in the anomalous component. The method of particle identification used in instruments to date means that this last point could not be confirmed. The method uses two measurements, the rate of energy loss dE/dx and the total energy, E . For a molecule these are proportional to $\sum_i Z_i^2$ and $\sum_i A_i$ respectively since each component atom has the same incident velocity (Z_i and A_i are the atomic number and mass of the i th component). Thus OH^+ closely resembles O^+ , and CO^+ will lie between Ne^+ and Na^+ in the data from these detectors. The most clearcut way of establishing that an individual event is a molecular ion involves the determination of dE/dx at several different energies covering the whole range of a stopping particle. For molecules consisting of different atoms, for example CO^+ , the different range of each component produces a distribution of dE/dx plotted against depth having two (or more) maxima. For the important case of molecules containing hydrogen such as OH^+ the range of the proton will often be sufficiently large to cause the event to be ignored unless special precautions are taken to recognise such events.

Recent high precision experiments on the ISEE and Voyager spacecraft⁹⁻¹¹ do not seem to have sufficient range discrimination to identify a molecular ion in this way. However, the high precision with which dE/dx and E are measured in these experiments should enable these molecular ions to be recognised as a group apparently having an anomalous mass, for example, CO^+ would appear to be ^{28}Ne and OH^+ would appear to be $^{17,25}\text{O}!$

If fluxes of molecular ions are reasonably high, a considerable number of different species could probably be detected. Their relative abundance would give valuable insight into processes associated with the injection of material and its acceleration to produce this particular low energy component of the cosmic radiation.

P. H. FOWLER
R. M. REDFERN
S. P. SWORDY

H. H. Wills Physics Laboratory,
University of Bristol,
Tyndall Avenue,
Bristol, UK

Received 12 March; accepted 11 May 1979.

1. Jokipii, J. R. *Rev. Geophys. Space Phys.* **9**, 27-87 (1971).
2. Garcia-Munoz, M., Mason, G. M. & Simpson, J. A. *Astrophys. J.* **202**, 265-275 (1975).
3. Hovestadt, D., Vallmer, O., Gloekler, G. & Fan, C. Y. *Astrophys. J.* **212**, 290-299 (1977).
4. Fisk, L. A., Kozlovski, B. & Ramaty, R. *Astrophys. J.* **190**, L35-L37 (1974).
5. Delsemme, A. H. *Space Sci. Rev.* **15**, 89-101 (1973).
6. Whipple, F. L. & Huebner, W. F. A. *Rev. Astr. Astrophys.* **14**, 143-172 (1976).
7. Axford, W. I. *Solar Wind* (eds Sonett, C. P., Coleman, P. J. & Wilcox, J. M.) 609 (NASA 1972).
8. Siscoe, G. L. & Mukherjee, N. R. *J. geophys. Res.* **77**, 6042-6051 (1972).
9. *IEEE Trans. on Geosci. Electronics*, GE-16/3, 151-280 (1978).
10. Krimigis, S. M. *et al. Space Sci. Rev.* **21**, 329-354 (1977).
11. Stone, E. C. *et al. Space Sci. Rev.* **21**, 355-376 (1977).

Increase of X-ray reflection intensities and profile widths at the low- to high- V_3O_5 phase transition state

An unexpected increase in peak intensity and profile half width has been observed in an X-ray diffraction experiment for many reflections from a crystal of V_3O_5 kept at its phase transition temperature. One explanation of the phenomena is that the 'mosaic structure' of the crystal is altered in the phase transition state, resulting in decreased extinction and increased profile widths. The results presented here are interesting spin off products of crystallographic studies of the mixed-valency oxide V_3O_5 . The semiconductor low- V_3O_5 transforms in a first order transition at 155°C into high- V_3O_5 , which is a 'poor metal'¹. The transformation is instantaneous and reversible and implies that an ordered distribution of V(III) and V(IV) atoms is replaced by an only partially ordered distribution, the changes in the atomic positions being fairly small ($\leq 0.1 \text{ \AA}$). These facts are based on the results of accurate X-ray crystal structure determinations of the two modifications at 25°C (ref. 2) and 185°C (ref. 3) respectively. We started diffraction measurements of V_3O_5 at or near the phase transition temperature t_T , using the good temperature-control facility of the available non-ambient-temperature equipment⁴.

We found that for many strong reflections the peak intensity increased drastically when the temperature passed t_T in either direction (see Fig. 1). Note that the intensity value for any particular strong reflection usually differs very little between the two modifications. The small difference between the crystal structures manifests itself by the presence of many very weak reflections hkl with $h+k+l=2n+1$ in low- V_3O_5 (space group $P2/c$) which are all systematically absent in high- V_3O_5 (space group $I2/c$). There is a general small decrease in the reflection intensities as the temperature of the crystal is increased from room temperature to just below t_T dependent on increased thermal motions of the atoms. However, the relative intensity decrease is considerably larger for reflections hkl with $h+k+l=2n+1$ than for the rest. Therefore a (minor) continuous structural change of low- V_3O_5 with temperature should precede

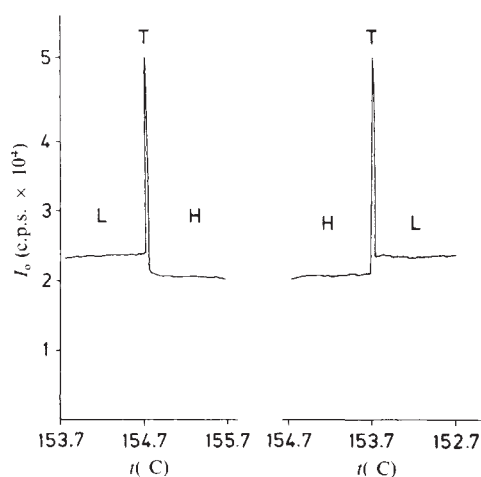


Fig. 1 Peak intensity of reflection 020 recorded with increasing and decreasing temperature respectively. Hysteresis effect is 1°C . A PAILRED X-ray single crystal diffractometer (inclination geometry) was used with $\text{MoK}\alpha$ radiation (graphite monochromator). The specimen crystal had an approximately cubic shape, the edge being about 0.17 mm . It was heated by hot nitrogen gas blown directly on to the crystal. The delivery and heating of the gas, and the temperature control were accomplished by a specially developed device⁴.

the discontinuous change at t_T . Thus, the difference around t_T between the two modifications of V_3O_5 is probably even smaller than stated above.

The unit cell dimensions of low- V_3O_5 at 25 °C are⁵ $a = 9.859(1)$, $b = 5.0416(5)$, $c = 6.991(1)$ Å and $\beta = 109.478(6)^\circ$. The changes from low- V_3O_5 at 150 °C to high- V_3O_5 at 160 °C are⁶ $\Delta a = -0.13\%$, $\Delta b = -0.18\%$ and $\Delta c = 0.13\%$. The close similarity between the unit cell dimensions of the two phases made the positions of the reflections, especially the low-angle ones, shift very little in crystal rotation angle ω at the transition, which was important for the initial observation of the effect.

One explanation of the phenomenon is that the rearrangement disorder in the transition state reduces the degree of perfection of the crystal with decreased extinction as one result. To test this hypothesis, many reflections with severe extinction in different directions of the reciprocal space were investigated. The intensity increase at t_T was always observed, at least for $y \leq 0.9$, y being the secondary extinction correction factor calculated according to Zachariasen⁷, and $y = I_{\text{obs}}/I_{\text{corr}}$ where I is the integrated intensity. Table 1 presents some of the studied reflections.

Although some of the strongest reflection intensities could be expected to suffer from counting losses no corrections were

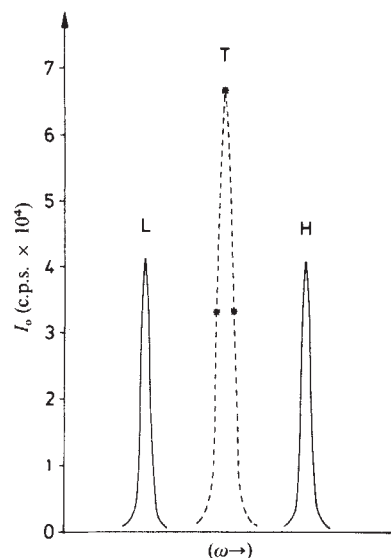


Fig. 2 Intensity profiles (ω -scan) for reflection -114 observed below (L) and above (H) t_T and constructed at (T) t_T from point observations (*). Experimental conditions as in Fig. 1.

Table 1 Secondary extinction correction (SEC), peak intensity and peak width: A comparison between low- V_3O_5 (L) at 153 °C, transition-state- V_3O_5 (T) and high- V_3O_5 (H) at 156 °C. For each reflection hkl first row means L, second T and third H

h	k	l	$\nu(^{\circ})$	$Y(^{\circ})$	SEC $y = \frac{I_{\text{obs}}}{I_{\text{corr}}}$	Peak intensity (c.p.s.)	Peak width (h) at half peak intensity ($^{\circ}$)	Mean increase in peak width (Δh), at $t_T(^{\circ})$
2	0	0	0.0	8.8		22,300	0.344	
						41,700	0.390	
					0.32	22,100	0.345	0.045
6	0	0	0.0	26.5		7,200	0.339	
						10,600	0.373	
					0.60	6,700	0.342	0.032
8	0	0	0.0	35.6		7,700	0.344	
						12,760	0.377	
					0.54	7,300	0.342	0.034
16	0	0	0.0	75.6		1,440	0.440	
						1,790	0.481	
					0.80	1,420	0.443	0.040
0	2	0	0.0	16.2		24,800	0.339	
						51,800	0.388	
					0.27	22,000	0.342	0.048
0	4	0	0.0	32.8		8,700	0.347	
						16,480	0.362	
					0.55	8,600	0.352	0.013
0	6	0	0.0	50.1		6,600	0.362	
						10,900	0.398	
					0.59	5,800	0.366	0.034
0	8	0	0.0	68.8		1,760	0.408	
						2,110	0.474	
					0.83	1,580	0.414	0.063
0	0	2	5.8	4.1		49,000	0.599	
						84,000	0.728	
					0.23	48,800	0.590	0.133
-1	1	4	11.7	9.1		41,400	0.612	
						66,900	0.878	
					0.26	41,000	0.607	0.269
-2	0	6	17.7	3.8		7,880	1.322	
						10,000	1.787	
					0.77	7,700	1.316	0.468
0	0	10	30.5	24.2		1,920	0.652	
						2,200	0.843	
					0.87	1,840	0.641	0.197

ν is the inclination angle of the detector axis, Y the angle between the detector axis and a horizontal plane containing the primary beam and the crystal rotation axis ($[[001]]$). The y -values actually refer to H but should be almost identical for L.

applied. However, this would mean that the real effect in these cases was even larger than the one observed.

If the crystal perfection is decreased in the phase transition state one should also observe a broadening of the width of the integrated intensity profile. Accordingly, comparative measurements of the half widths of these profiles for the reflections included in Table 1 were made just below t_T , at t_T as well as just above it. The measurements at the phase transition temperature were difficult to perform, and the time available for each of them usually was between 5 and 10 s. However, as the actual reflections had strong intensities, using the pen recorder as the intensity indicator was successful. In all cases investigated, the reflection profile more or less broadened at t_T (see Fig. 2 and Table 1). This quantity is a multiple convolution of a number of experimental components of which all but one or possibly two crystal mosaicity components, isotropy being assumed, should be fairly constant for each reflection at and around t_T . An attempt will be made to correlate the observed half width increases at t_T with the theoretical expressions for the mosaic angle distribution half width and the mosaic particle size component half width, which have both been derived by Ito⁸. However, a deconvolution analysis is necessary for this which is still in progress.

Similar effects of other comparable phase transitions do not seem to have been discussed previously. However, the transient intensity increase of a strong reflection at the α - to β -quartz transition which is shown in ref. 9 may be another example.

STIG ÅSBRINK
SAM-HYO HONG

Department of Inorganic Chemistry,
Arrhenius Laboratory,
University of Stockholm,
S-106 91 Stockholm, Sweden

Received 5 January; accepted 20 April 1979.

1. Chudnovskii, F. A., Terukov, E. I. & Khomskii, D. I. *Solid State Commun.* **25**, 573–577 (1978).
2. Åsbrink, S. *Acta crystallogr.* (submitted).
3. Hong, S.-H. & Åsbrink, S. *Acta crystallogr.* (in the press).
4. Hong, S.-H. & Åsbrink, S. *J. appl. Crystallogr.* (in the press).
5. Åsbrink, S. *Mat. Res. Bull.* **10**, 861–863 (1975).
6. Brückner, W., Moldenhauer, W., Thuss, B. & Försterling, G. *Phys. Stat. Sol. (a)* **35**, K49–K51 (1976).
7. Zachariasen, W. H. *Acta crystallogr.* **23**, 558–564 (1967).
8. Ito, T. *RIKEN Rep.* **47**, 47–60 (1971).
9. Young, R. A. *J. Scient. Instrum.* **43**, 449–453 (1966).

Measurement of the neutron spectra from beam-heated PLT plasmas

MEASUREMENTS of the neutron spectra from the PLT tokamak plasma are reported here which provide information on the centre-of-mass velocity of the reacting deuteron pairs. The observation in PLT of no significant directed velocity yields additional evidence that the neutrons are produced by a thermonuclear process. Previous spectral measurements on other fusion devices have often determined that the neutron emission was non-thermonuclear and resulted from a small 'tail' component of energetic ions¹⁻³. The neutron spectra indicated a directed velocity of the reacting deuterons and thus non-thermonuclear neutron emission if the injected neutral beam ions were deuterium. In this condition, the neutron spectra were consistent with the expected beam-induced $d(d,n)^3\text{He}$ reactions. However, if the heating beam used hydrogen, then the neutron spectra are consistent with a thermonuclear neutron emission resulting from the beam heating of the bulk deuterons in the plasma.

The importance of a thermonuclear neutron emission is that it allows the neutron emission measurement to be used to deduce ion temperature⁴. The ion temperature deduced by the magnitude of the neutron emission has been consistent with other PLT ion temperature diagnostic^{5,6} such as charge exchange spectral observations. Although the charge exchange spectrum indicated a maxwellian ion population up to $6T_i$, the observation had been made only perpendicularly to the direction of the tangential neutral beam injection. Thus, we use the neutron spectral measurements to provide information on the plasma ion population in the direction parallel to the neutral beam injection where non-thermonuclear emission might be induced by the

presence of trace amounts of deuterium in the hydrogen beams or by momentum transfer from the beams to the background plasma. To observe such asymmetries, a neutron spectrometer was collimated (Fig. 1) to look towards co-injected ions and comparisons were made between co-only injection and counter-only injection. In this manner, any non-thermonuclear emission induced by the beams would be observed from the shift of the mean neutron energy caused by the directed centre-of-mass velocity of the reacting deuteron pairs. Thus, the PLT neutron spectra experiment was carried out in the manner of the classic neutron spectra experiments on Zeta¹. All the PLT spectra have about as many counts as in those Zeta spectra, although for PLT the spectra were obtained with time resolution and with higher energy resolution.

The PLT tokamak⁷, and its operation⁵ with neutral beams have been described elsewhere. For these experiments, the plasma had a minor radius of 0.40 m as defined by carbon limiters, a major radius of 1.34 m, a toroidal magnetic field of 3.2 T, a plasma current of 0.4 MA, and titanium gettered vacuum walls. Four tangential neutral beams which can collectively deliver up to 2.5 MW of power to the plasma were used. Two of the beams (co) were directed parallel to the plasma current, and two (counter) were directed antiparallel to the plasma current.

About 200 discharges were run with hydrogen neutral beam injection into a deuterium plasma ($\text{H}^0 \rightarrow \text{D}^+$) and about 15 discharges were run with deuterium neutral beam injection into a deuterium plasma ($\text{D}^0 \rightarrow \text{D}^+$). In each set of discharges, the co-injectors were used early in the discharge and the counter-injectors were applied later in the same discharge after the co-injectors were turned off and the neutron emission from them had decayed. The neutron emission (Fig. 2) was about $7 \times 10^{10} \text{ n s}^{-1}$ during the hydrogen injection ($\text{H}^0 \rightarrow \text{D}^+$) and about 10^{12} n s^{-1} during the deuterium injection ($\text{D}^0 \rightarrow \text{D}^+$). For the $\text{D}^0 \rightarrow \text{D}^+$ cases, only one beam at reduced voltage was used to

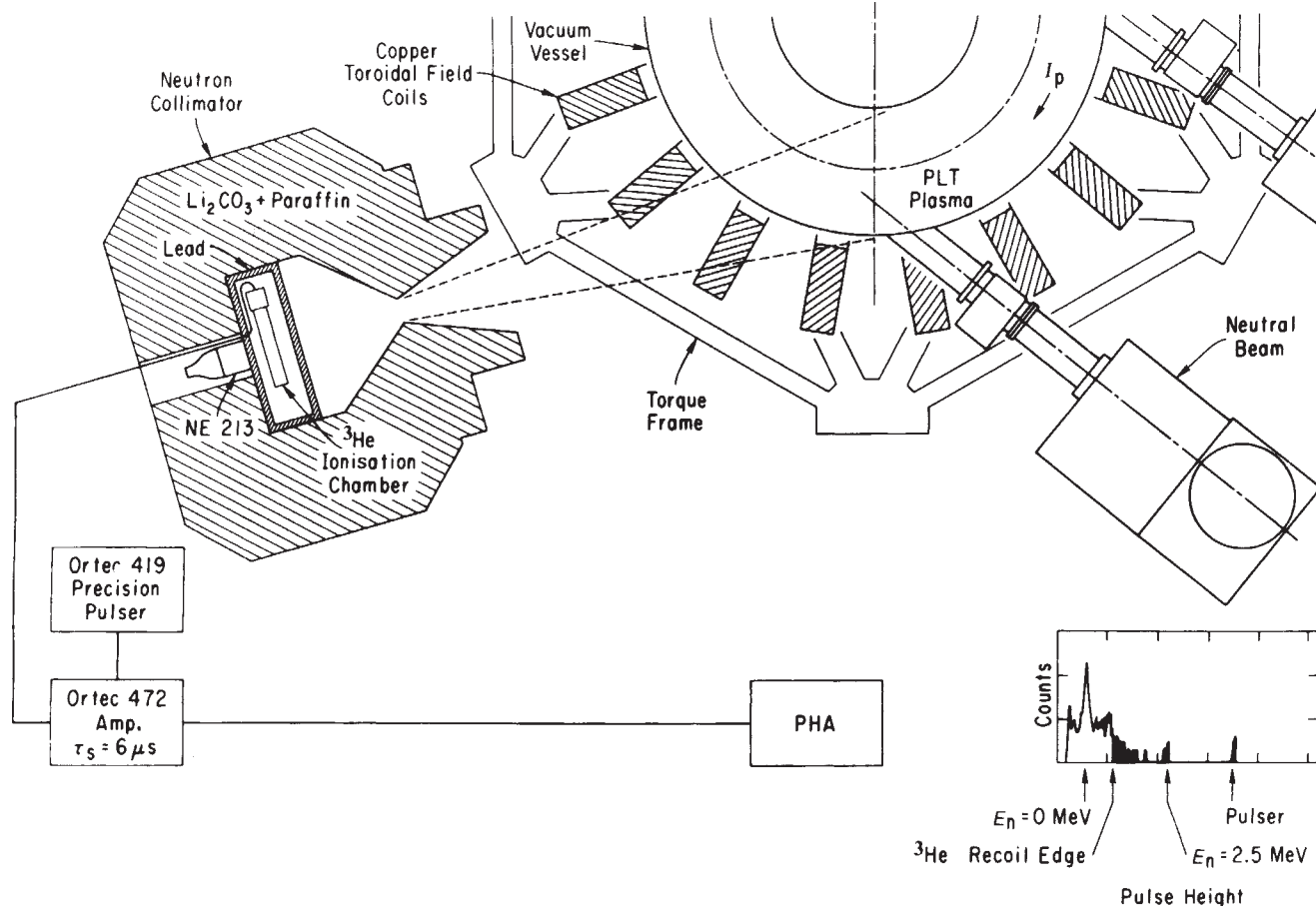


Fig. 1 The ^3He ionisation chamber is surrounded by two inches of lead to reduce X-ray induced counts and is collimated by a 7-ton paraffin plus Li_2CO_3 neutron collimator to view tangentially to the PLT minor axis.

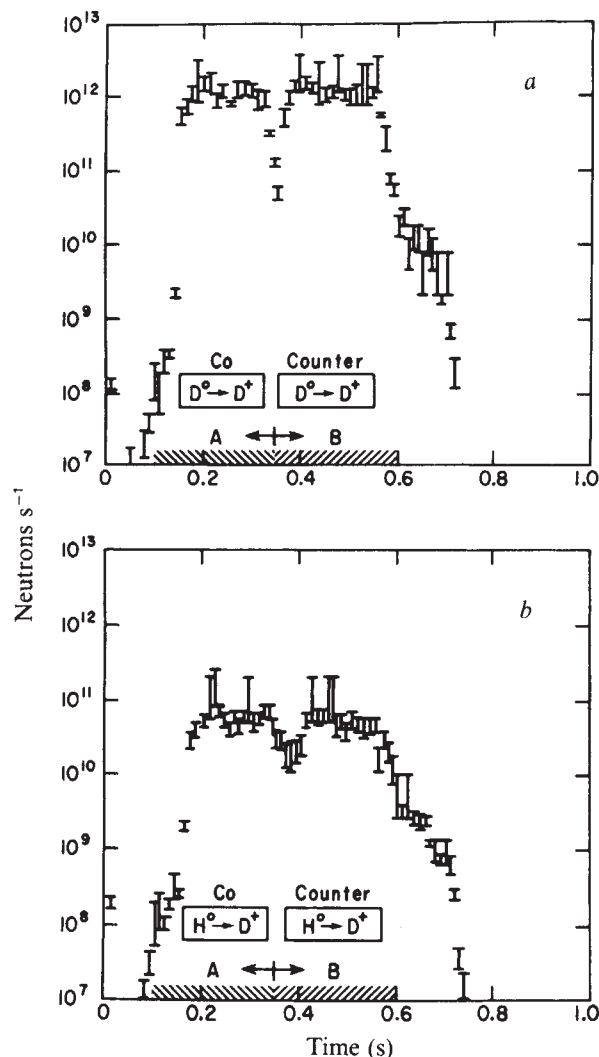


Fig. 2 The neutron emission during (a) $D^0 \rightarrow D^+$ (about 0.3 MW in each direction) and (b) $H^0 \rightarrow D^+$ (about 0.9 MW in each direction). The detector counting intervals were made to coincide with A, co-only injection; and B, counter-only injection. The ion temperature deduced from the magnitude of the neutron emission ($H^0 \rightarrow D^+$) is about 2 keV.

keep the neutron emission low and thus prevent the neutron spectrometer from saturating.

The neutron spectra were recorded by a ^3He ionisation chamber (Fig. 1) and the counts were recorded separately for two different time zones during the discharge. The first time zone encompassed and recorded the co-only injection. The second time zone encompassed and recorded the counter-only injection. The early and end stages of the discharge were excluded to remove any possibility of X-ray noise induced by runaway electrons. All the discharges reported here were quite free of runaway electrons.

The neutron collimator, which is similar to one used by Glasgow *et al.*⁸, is composed of equal weights of paraffin and Li_2CO_3 to a linear dimension of about nine absorption lengths for 2.5 MeV neutrons. The paraffin thermalises the neutrons incident on the collimator from directions other than the observation volume, and the ^6Li in the natural lithium of the Li_2CO_3 non-radiatively absorbs the thermal neutrons. Tests of the collimator with the observation window plugged, indicate that $<10^{-3}$ of the total spectrometer counts occur from neutrons incident on the ^3He detector from unobserved regions. The complex PLT geometry (Fig. 1) enables many neutron scattering centres to be within the line of sight, such as the massive copper coils, diagnostic equipment, and the concrete wall on the far side of PLT. These scattering centres have an unknown influence on

the recorded neutron spectra and should introduce low energy neutrons into the spectra.

The ^3He ionisation chamber^{9,10} collects the charge produced by the proton and triton in the ^3He (n,p)t reaction. The detector saturates at total count rates of about 10^4 c.p.s., but even near saturation, only a few c.p.s. are recorded in the region near 2.5 MeV. For each ^3He (n,p)t count near 2.5 MeV, there are many more ^3He recoil counts, and ^3He (n,p)t counts for scattered neutrons which have reached epithermal energies (spectrum in Fig. 1). The absolute energy scale was derived from the Q of the ^3He reaction (0.76 MeV) and the linear response of the spectrometer. This procedure could result in the energy scale of the set of spectra being in error by ± 0.05 MeV in the region near 2.5 MeV. The spectrometer was calibrated using accelerator induced mono-energetic neutrons and has an energy resolution of about 0.055 MeV FWHM, or standard deviation of a single count. The ability of the spectrometer to determine relative mean energies should be limited only by statistics.

The $D^0 \rightarrow D^+$ spectra (Fig. 3) indicate an energy shift of 0.26 MeV between co-only and counter-only injection, establishing the ability of the ^3He detector to make noticeable neutron energy shift measurements in the PLT conditions. Each of these $D^0 \rightarrow D^+$ spectra can be reasonably explained by the expected beam-induced reactions between the energetic injected ions and the background thermal ions. The PLT $D^0 \rightarrow D^+$ spectra were compared to $Z_{\text{eff}} = 5$, $T_i = 1.5$ keV Monte Carlo

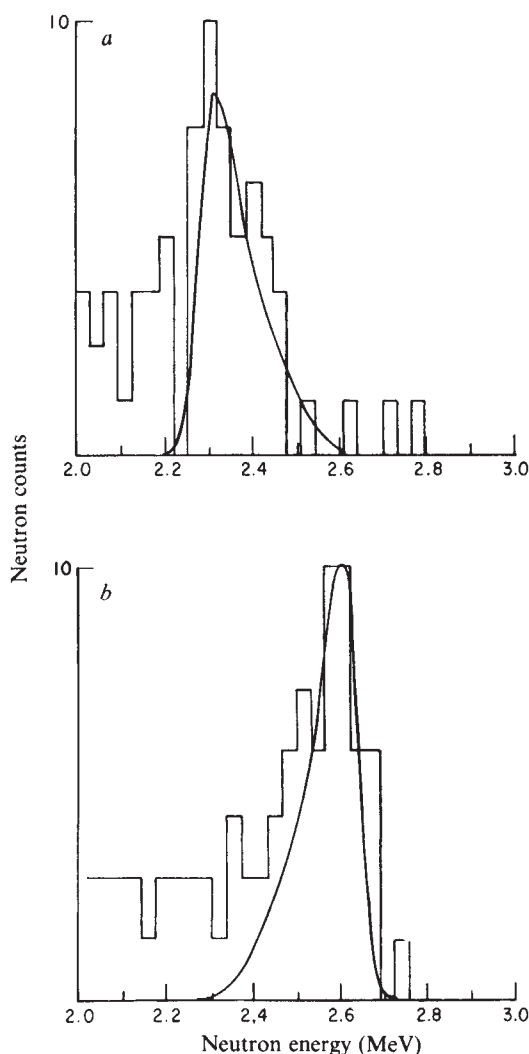


Fig. 3 The measured $D^0 \rightarrow D^+$ neutron spectra. The counter-injection voltage was 30 kV (a). The co-injection voltage was 26 kV (b).

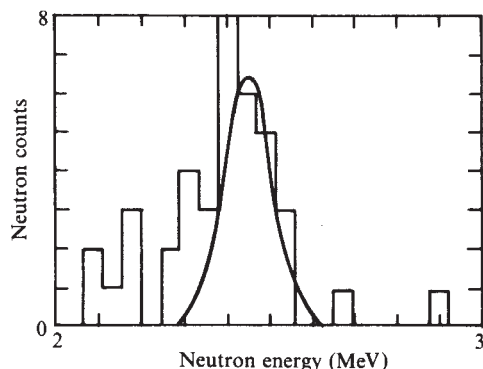


Fig. 4 The measured $H^0 \rightarrow D^+$ neutron spectra including all the co-only and counter-only counts.

code predictions¹¹ for the collimated neutron spectra (Fig. 3). The co-injection case indicated a 60% χ^2 probability of being produced by beam-induced reactions while the counter-injection case indicated a 55% χ^2 probability of being produced by beam-induced reactions. The χ^2 test was performed only for that part of the calculated spectrum which was >10% of the peak.

The $H^0 \rightarrow D^+$ spectra indicate no statistically significant energy shift between the co-only and counter-only injection cases, as is expected for thermonuclear neutron emission. The predicted mean energy should be 2.45 MeV, while experimentally, the 11 co-only counts have a mean energy of 2.41 ± 0.02 MeV and the 19 counter-only counts have a mean energy of 2.42 ± 0.01 MeV. The downward energy shift from 2.45 MeV probably results from the energy scale calibration. The expected line width should be 0.12 MeV for the 2 keV ion temperature plasma measured by a detector with 0.055 MeV resolution. Adding all the experimental $H^0 \rightarrow D^+$ counts together to improve the statistics (Fig. 4) results in a line width of 0.13 MeV. The observed $H^0 \rightarrow D^+$ neutron spectrum indicated a 50% χ^2 probability of being produced by thermonuclear reactions. One degree of freedom was used in the χ^2 test to normalise the mean of the spectra and the fit was applied only for that part of the calculated spectrum >10% of the peak.

In conclusion, no measurable centre-of-mass velocity of the reacting deuteron pairs was induced by hydrogen neutral beam injection. Furthermore, the shape of the experimental neutron emission line was consistent with a thermonuclear neutron emission at the ion temperature predicted by the magnitude of the neutron emission. These two observations are consistent with a thermonuclear origin of the PLT neutrons during hydrogen neutral beam heating of a deuterium plasma.

We thank J. Boychuk, W. Mycock, M. Perron and the PLT crew for experimental help. The neutron spectrometer was tested at RTNSI with the help of D. Slaughter. This work was supported by the US Department of Energy, Contract no. EY-76-C-02-3073. L.S. is from Exxon Nuclear Co. and R. S. from Westinghouse Corporation.

J. D. STRACHAN

P. COLESTOCK

H. EUBANK

L. GRISHAM

J. HOVEY

G. SCHILLING

L. STEWART

W. STODIEK

R. STOOKSBERRY

K. M. YOUNG

Plasma Physics Laboratory,
Princeton University,
Princeton, New Jersey 08544

Received 23 February; accepted 4 May 1979.

1. Rose, B., Taylor, A. E. & Wild, E. *Nature* **181**, 1630–1632 (1958).
2. Combe, R. A. & Ward, B. A. *Plasma Phys.* **5**, 273–283 (1963).
3. Pecorella, F. *et al. Phys. Fluids* **20**, 675–682 (1977).
4. Artsimovich, L. A. *et al. Sov. Phys. JETP* **34**, 306–309 (1972).
5. Eubank, H. *et al. Plasma Physics and Controlled Nuclear Fusion Research* (paper IAEA-CN-37-C-3, Innsbruck, 1978).
6. Brusati, M. *et al. Nucl. Fus.* **18**, 1205–1216 (1978).
7. Grove, D. *et al. Plasma Physics and Controlled Nuclear Fusion Research* **1**, 21–31 (1977).
8. Glasgow, D. *et al. Nucl. Instrum. Meth.* **114**, 521–534 (1974).
9. Cuttler, J. M., Shalev, S. & Dagan, Y. *Trans. Am. nucl. Soc.* **12**, 63 (1969).
10. Franz, H. *et al. Nucl. Instrum. Meth.* **144**, 253 (1977).
11. Towner, H. H. & Jassby, D. L. *Trans. Am. nucl. Soc.* **22**, 74–75 (1975).

Deposition and accumulation of plutonium isotopes in Antarctica

DATA on the deposition of plutonium isotopes are presented here from the atmosphere at Dome C ($123^{\circ}10' E$, $74^{\circ}39' S$; 3,214 m elevation) on the high Antarctic plateau. Plutonium isotopes are among the anthropogenic chemicals that have become global contaminants^{1–3} and it is, therefore, important to gain a historical perspective to their worldwide dispersion. The analysis of successive layers of permanent snow fields permits the determination of both present and historical fluxes of anthropogenic chemicals and other contaminants that are dispersed through the atmosphere^{3–6}. Dome C is an ideal site for such studies: annual precipitation at Dome C is of the order of $3.7 \text{ g H}_2\text{O cm}^{-2}$ (ref. 7), considerably lower than the mean annual deposition of $15.5 \text{ g H}_2\text{O cm}^{-2}$ over the entire continent⁸. The mean annual temperature is -53.5°C ⁷, with summer temperatures remaining well below freezing, precluding vertical percolation through successive layers and reducing potential losses from volatilisation.

A 5 m pit was dug at Dome C in January 1977. Snow in layers of approximately 10 cm was sampled and the meltwater was transferred into 20-l, polyethylene cubitainers. They were then acidified, equilibrated, and the isotopes subsequently co-precipitated with aluminium phosphate following the technique of Koide and Bruland⁹. Each sample consisted of a total of 100 litres. The precipitates in the original cubitainers were returned to the US. They were dissolved in 9 M HCl and the Pu isotopes isolated following the procedures of Talvitie¹⁰ and Wong.¹¹ An aliquot was removed for ^{210}Po determination⁹. Samples were counted within 2 weeks after separation by α spectrometry; ^{208}Po and ^{242}Pu were used as yield tracers.

Figure 1 shows the activities of $^{239+240}\text{Pu}$ and ^{238}Pu versus depth of snow. Assignment of the 1955 time horizon to 214 cm, at the beginning of the rise in ^{239}Pu activity, was based on previous work at Dome C by Lorius⁷ and studies undertaken elsewhere in Antarctica^{12–14} which have shown a sharp increase in ^{90}Sr and gross β activity in strata deposited in early 1955. This assignment is consistent with our ^{210}Pb results. Assuming constant sedimentation for the depth interval, 0–250 cm, ^{210}Pb activity at any depth z , A_z , can be expressed by $A_z = A_0 \exp(-\beta z)$. Our data yielded an A_0 of $142 \text{ d.p.m. per } 100 \text{ litres}$ and a β of -0.00308 cm^{-1} , where $r = 0.90$ for the best-fit curve. Thus, a time interval of $21.2 \pm 2.4 \text{ yr}$ (1σ) to this 214 cm depth was calculated. Our assignment of the 1965 time horizon interval was derived by interpolation and estimates of maximum error due to compaction¹⁵.

The profile of $^{239+240}\text{Pu}$ deposition (Fig. 1a) is unlike the profile observed in Greenland⁵ in that the majority of the total fallout occurred before 1960, whereas in Greenland the greater portion of the deposition occurred after 1960. Significant quantities of $^{239+240}\text{Pu}$ began to appear in the snow layer corresponding to the estimated date of 1955. The ensuing peak's size and timing suggest that a significant portion of the $^{239+240}\text{Pu}$ could have been contributed by the US Castle test, Bravo, on 28 February 1954, at Bikini ($165^{\circ} E$, $11^{\circ} N$). This 15 megatonne detonation derived a major portion of its energy from the

fusion-induced fission of ^{238}U (ref. 16). Large quantities of ^{239}Pu were probably produced from ^{238}U by neutron capture. Previous studies¹²⁻¹⁴ have shown that significant stratospheric fallout from this test began to arrive in the Antarctic 12 ± 2 months later. Further weapons tests at Bikini and Eniwetok such as the Hardtack series during 1958 might have contributed additional $^{239+240}\text{Pu}$ to this major peak. ^{90}Sr attributed to these tests has been observed at several Antarctic sites with an arrival date of early 1959¹⁵. The smaller amounts of $^{239+240}\text{Pu}$ deposited at Dome C after 1960 suggest that the tests conducted in the early 1960s, primarily in the Northern Hemisphere, resulted in a relatively low level of stratospheric transfer to the high latitudes of the Southern Hemisphere.

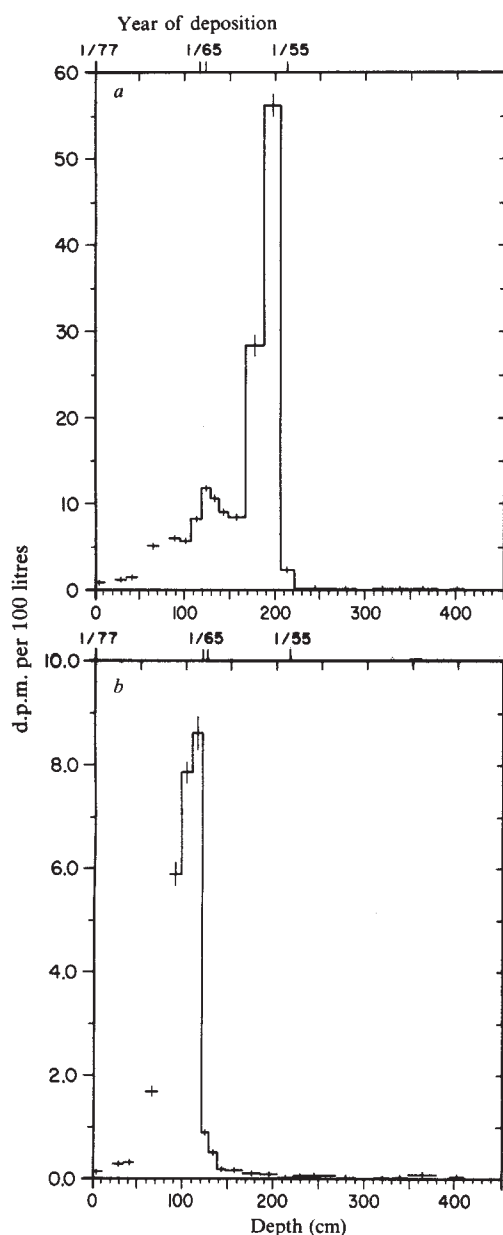


Fig. 1 $^{239+240}\text{Pu}$ (a) and ^{238}Pu (b) activities at Dome C, Antarctica. Horizontal lines depict the depth of each layer sampled. Vertical lines within each sample represent the uncertainty due to counting statistics (1σ). Plutonium sample counting times ranged from 2 to 9 d. The minimum detectable amount of ^{238}Pu and $^{239+240}\text{Pu}$ was found to be of the order of 0.03 d.p.m. per 100 litres. Samples for the depth intervals, 11.9–23.9 cm, 48.0–59.9 cm, and 71.9–83.8 cm were not obtained.

Figure 1b shows the snow record of ^{238}Pu at Dome C. The April 1964 burnup of a satellite (SNAP-9A) containing 17 kCi of ^{238}Pu in the Southern stratosphere¹ is recorded as a rapid activity increase in the 1965–66 stratum. The deposition of ^{238}Pu closely parallels the changes in the stratospheric inventory of SNAP ^{238}Pu over the period 1964–69 as shown by Harley¹⁷. The activity ratio of $^{238}\text{Pu}/^{239+240}\text{Pu}$ also rises sharply, reaching a maximum of 150% in the layer assigned to approximately 1967 (the stratum following the peak activity). Harley¹⁷ reported comparable high activity ratios in Southern Hemisphere surface air samples in 1967 and 1968.

During the period of peak fallout of $^{239+240}\text{Pu}$ (1955–59), levels of ^{238}Pu activity were markedly low (^{238}Pu values of 0.10 ± 0.02 and 0.12 ± 0.03 d.p.m. per 100 litres compared to the respective $^{239+240}\text{Pu}$ values of 56.2 ± 1.3 and 28.4 ± 1.2 d.p.m. per 100 litres). Harley¹⁷ suggested that a $^{238}\text{Pu}/^{239+240}\text{Pu}$ ratio of 3% might be representative of weapons fallout. If this was the case, then a significant ^{238}Pu peak should have been associated with the maximum fallout of $^{239+240}\text{Pu}$. However, this was not observed, and in fact, the activity ratio was of the order of 0.3%, an order of magnitude lower. Harley¹⁷ pointed out that the value of 3% was not well established since good α spectrometry was not available on their samples before SNAP-9A. Thus, this ratio may be doubtful as a widely applicable value. This discrepancy warrants further investigation.

To calculate the fluxes of $^{239+240}\text{Pu}$ and of ^{238}Pu , a mean density of 0.352 g cm^{-3} , derived from the data of Lorius⁷ was used. The flux of $^{239+240}\text{Pu}$ was calculated by integrating over the depth interval 0–222.5 cm. This value was found to be $90.4 \text{ d.p.m. m}^{-2}$, while the peak of $19.2 \text{ d.p.m. m}^{-2} \text{ yr}^{-1}$ occurred during 1956–57. Similarly, after decay correction the flux of ^{238}Pu was determined to be $14.4 \text{ d.p.m. m}^{-2}$. The peak flux of $3.22 \text{ d.p.m. m}^{-2} \text{ yr}^{-1}$ occurred around 1965–66. Hardy *et al.*¹ using soil data have estimated cumulative worldwide fallout of plutonium isotopes as a function of latitude. Their estimates for $70\text{--}80^\circ \text{S}$, based on extrapolation from Southern Hemisphere samples, were $0.03 \pm 0.01 \text{ mCi km}^{-2}$ of $^{239+240}\text{Pu}$ and $0.008 \pm 0.005 \text{ mCi km}^{-2}$ of ^{238}Pu derived from SNAP-9A. Our results provide corresponding values of 0.04 and 0.006, respectively, in good agreement with the estimates cited above.

The profiles of Fig. 1 suggest that almost all of the plutonium isotopes injected into the atmosphere have now been removed. The activities of both $^{239+240}\text{Pu}$ and ^{238}Pu deposited during 1976 at Dome C were only 1.4% of the activities deposited during the respective periods of maximum fallout.

G. A. CUTTER
K. W. BRULAND

Center for Coastal Marine Studies,
University of California,
Santa Cruz, California 95064

R. W. RISEBROUGH

University of California,
Bodega Marine Laboratory,
Bodega Bay, California 94923

Received 21 February; accepted 23 April 1979.

- Hardy, E. P., Krey, P. W. & Volchok, H. L. *Nature* **241**, 444–445 (1973).
- Bowen, V. T., Wong, K. M. & Noshkin, V. E. *J. mar. Res.* **29**, 1–10 (1971).
- Koide, M., Goldberg, E. D., Herron, M. M. & Langway, C., Jr *Nature* **269**, 137–139 (1977).
- Murozumi, M., Chow, T. J. & Patterson, C. *Geochim. cosmochim. Acta* **33**, 1247–1294 (1969).
- Weiss, H. V., Koide, M. & Goldberg, E. D. *Science* **174**, 692–694 (1971).
- Cragin, J. H., Herron, M. M. & Langway, C. C., Jr *U.S. Army Cold Regions Research and Engineering Laboratory (USACRREL) Research Report 341* (Hanover, New Hampshire, 1975).
- Lorius, C. *Ant. J. of U.S.* **10**, 159 (1975).
- Bull, C. in *Research in the Antarctic* (American Association for the Advancement of Science, Washington, 1971).
- Koide, M. & Bruland, K. W. *Analyt. chim. Acta* **75**, 1–19 (1975).
- Talvitie, N. A. *Analyt. Chem.* **43**, 1827–1830 (1971).
- Wong, K. M. *Analyt. chim. Acta* **56**, 355–364 (1971).
- Picciotto, E. & Wigain, S. *J. geophys. Res.* **68**, 5965–5972 (1963).
- Wigain, S., Picciotto, E. & DeBreuck, W. *J. geophys. Res.* **70**, 6023–6031 (1965).
- Woodward, R. N. *Nature* **204**, 1291 (1964).
- Crozaz, G., Picciotto, E. & DeBreuck, W. *J. geophys. Res.* **69**, 2597–2604 (1964).
- Rotblat, J. *Bull. atom Sci.* **11**, 171–177 (1955).
- Harley, J. H. *Proc. Envir. Plutonium Symp.* LA-4756 (1971).

Climate and thermodynamic systems of maximum dissipation

THE Earth-atmosphere is a classic example of a closed, dissipative and nonlinear thermodynamic system which is subject to both regular and irregular impulses causing significant departure from steady state. It is closed because it exchanges energy (solar and thermal radiant energy) but not mass with its environment. It is dissipative because the net input of radiant energy occurs mainly in regions of high temperature towards the Equator and the net output occurs mainly in regions of low temperature towards the poles. It is nonlinear basically because of the multiplicity of internal feedbacks and because of the importance of advective processes. It has steady-state character in the sense that the annual mean radiant energy input is very close to the annual mean output, and parameters such as the annual mean temperature do not vary significantly from one period to another. The regular seasonal variation in solar position ensures significant departure from the steady state so defined, and there are also significant irregular departures arising (for instance) from variations of solar input and IR output caused by variations in the amount and distribution of cloud. Recently I have shown¹ that the overall Earth-atmosphere climate system seems to have adopted a format whereby the total thermodynamic dissipation associated with the horizontal energy flows in the atmosphere and ocean is a maximum. 'Format' in this context refers to the annual average geographic distribution of cloud, surface temperature and the horizontal energy flows. The practical significance of this is that, if one could accept it as a general principle governing climate behaviour, one could use it directly as a means of *a priori* prediction of climate and climate change without needing detailed analysis of the internal workings of the system. I could not explain previously why the Earth-atmosphere system should be so constrained. This note points out that the Earth-atmosphere has characteristics such that it might be expected to obey such a constraint. Furthermore, these characteristics are sufficiently general that the same principle of selection of steady-state mode of maximum dissipation may apply to a broad class of non-linear systems.

A fair thermodynamic representation of the global climate system is as depicted by the solid lines in Fig. 1. There is an input of solar energy I , most of which is converted directly into various forms of internal energy IE and is returned to space in the form of thermal IR radiation. Some of the input appears as available potential energy P , is converted to kinetic energy K , and is in turn dissipated by friction into internal energy before being radiated back to space. At steady state the rates of creation of

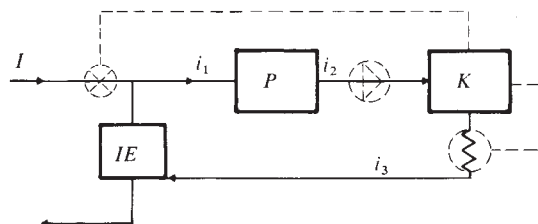


Fig. 1 The total Earth-atmosphere system of: energy transfers (thin solid lines); reservoirs of internal, available potential, and kinetic energy (IE , P and K respectively); feedback controls from kinetic energy (dashed lines); and an indication of the non-linearity of the rate of potential energy conversion (dashed diode symbol). Adequate data are not yet available for the oceanic part of the total system, but for the atmospheric component the annual average quasi-steady state values are: $i_1 = i_2 = i_3 \approx 2.3 \text{ W m}^{-2}$; $K \approx 1.5 \times 10^6 \text{ J m}^{-2}$ with the ratio $P:K \approx 10:1$ (refs 2, 7).

available potential energy, conversion of available potential to kinetic energy and the dissipation of kinetic energy (i_1 , i_2 and i_3 respectively) are equal.

The dashed components of Fig. 1 concern two characteristics of importance to the present argument. First, there are feedbacks such that both the rate of energy input (I or i_1) and the rate of dissipation i_3 are functions of, or can in principle be related to, the total kinetic energy. (One of the prime controls on the rate of energy input is, for instance, cloud cover; and cloud cover is a complex function of the amount and distribution of the large and small-scale motions of the atmosphere and ocean, of quantities which determine the motion, or of quantities which are determined by the motion.) Second, the rate of conversion i_2 of potential to kinetic energy is a highly non-linear and increasing function of potential energy. Some examinations of the process effectively envisage an $i_2 = (P - P_c)^n$ relation where P_c is some critical value of potential energy below which the rate of conversion is very small⁴ and $n > 1$.

Let i_1 and i_3 be functions of kinetic energy as in Fig. 2. The actual shape of the curves is irrelevant. The important characteristics are: (1) There is the possibility of a number (and in the case of the Earth-atmosphere system, a large number) of steady state positions where the rate of energy input equals the rate of output (that is $i_1 = i_3$); and (2) the dissipation rate i_3 is an increasing function of kinetic energy.

The real system is sufficiently complex in its detailed interactions for there to be considerable natural variability in, for instance, cloud cover and the rate of energy input. Therefore there is variability about whatever is the steady state determined by the current average value of K . The extent and character of this variability is unknown even for the quasi-steady state which the Earth-atmosphere system presently occupies. However, consider at some stage that the system was in one of the mid-range stable regimes (A of Fig. 2 say). The input I or i_1 will be subject to natural variability which may or may not be random and is not necessarily a function of K , but may occasionally be large enough to boost the system to another stable regime. If the variability is random, presumably an impulse of $+\Delta Q$ in i_1 is as likely to occur as an impulse of $-\Delta Q$, and in the absence of other information the system is as likely to change to a new steady state of lower average K as to one of higher average K .

However, the non-linearity of the conversion of P to K ensures that for large values of ΔQ (short term changes of i_1 away from its steady state value where they are not necessarily an immediate function of K)

$$\frac{\Delta K^+}{\Delta Q} > \frac{\Delta K^-}{-\Delta Q} \quad (1)$$

That is, a positive impulse ΔQ in i_1 is more likely to change the system to a steady state of higher average K (and dissipation) than is a negative impulse in i_1 likely to change the system to a steady state of lower average K . This is obvious (see legend to Fig. 2) but is not easy to prove in a rigorous manner. A complete analysis would require a fairly sophisticated application of control theory. (It would require knowledge of the exact form of the random (?) impulses, and some specific knowledge of at least the statistical distribution of steady states as a function of K . Such information is not yet available for the Earth-atmosphere system.) Suffice it to say that first order treatment leads to the relationship

$$\frac{\Delta K^+}{\Delta Q} - \frac{\Delta K^-}{-\Delta Q} \approx \frac{b\epsilon}{1 + 2cZ + c\epsilon + c^2 + c^2Z^2 + c^2Z\epsilon} \quad (2)$$

where, if Δt is the short time period over which the impulse is applied and a is the average slope of the i_3 versus K relationship and ΔP or $-\Delta P$ are the net responses in available potential energy to ΔQ or $-\Delta Q$ respectively, then

$$b = \frac{1}{2}\Delta t / (1 + \frac{1}{2}a\Delta t); c = \frac{1}{2}\Delta t; Z = \Delta i_2 / (-\Delta P); \text{ and } \epsilon = \Delta i_2 / \Delta P - \Delta i_2 / (-\Delta P)$$

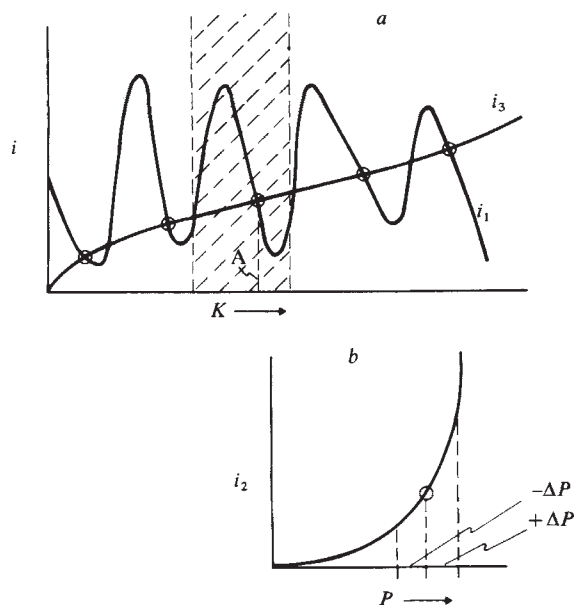


Fig. 2 *a*, The functions relating average rate of energy input i_1 and the rate of dissipation i_3 to the average total kinetic energy K . The circled points are possible stable steady states. The uncircled points where $i_1 = i_3$ are unstable steady states because of the positive feedback associated with slight displacements of K . The hatched area roughly indicates the extent of the stable regime about one such steady state A (that is between the adjacent unstable steady states). The central question is why short-term random variability of i_1 about the average value determined by K should (if the impulses are large enough) preferentially boost the system to steady states of higher average K and ultimately to maximum i_3 or dissipation. The basic reason for the system under discussion is super-linearity in the rate of conversion i_2 of available potential energy P (see *b*). Sub-linearity would tend to ensure movement to steady states of lower average K . The intuitive argument (see text) is as follows: $+\Delta Q$ on $i_1 \rightarrow$ growth of $P \rightarrow$ much increased $i_2 \rightarrow$ large growth of K . $-\Delta Q$ on $i_1 \rightarrow$ decay of $P \rightarrow$ slightly decreased $i_2 \rightarrow$ slight decay of K .

Note that the changes Δi_2 in Z and ε refer to the result of changes ΔP and $-\Delta P$ away from the steady state value of P . Thus, if one accepts $\Delta K^+ / \Delta Q - \Delta K^- / (-\Delta Q)$ as a measure of the 'strength' S of the tendency under random forcing to move to steady states of higher dissipation, then large S is ensured among other things by greater positive non-linearity ε in the rate of conversion of potential to kinetic energy. It is ensured also by smaller slope of the i_3 versus K relation.

Provided (1) The rate of conversion of P to K is a super-linear increasing function of P ; (2) the rate of dissipation is an increasing function of K ; (3) the system has several possible stable steady states; and (4) there is random variability in the rate of energy input which is, or was at some time, of sufficient magnitude to boost the system from one steady state to another, then it can be expected that the system will ultimately occupy that steady-state regime which has maximum dissipation. As far as the Earth-atmosphere is concerned all the conditions are acceptable, but are difficult to prove absolutely. In particular, it has yet to be proved that the dynamical constraints on atmospheric and oceanic motions allow the existence of more than one of the steady states otherwise allowable on energy balance considerations alone. Furthermore, the possibility must be considered that the distribution of possible steady states as a function of average kinetic energy is so extraordinary that the mechanism envisaged here is effectively negated.

Note that 'random variability of input' can be occasioned by processes internal to the system. It is not necessary to assume changes in some external parameter such as the solar constant.

In general terms the above limitations are sufficiently broad and non-restrictive to suggest that the overall concept may be applicable to a large class of non-linear thermodynamic systems

where there is 'series' conversion of energy from one form to another. In the Earth-atmosphere example, the non-linearity of the conversion of P to K is particularly obvious (ε of equation (2) is large) and presumably once the system is in its ultimate format there is no tendency to return to a regime of lower dissipation. However, one can imagine systems where the non-linearity is weak and the ultimate behaviour would be no more than a tendency to spend more time in regimes of higher dissipation. Similarly a system of strongly sub-linear energy conversion (negative ε) would be expected to move towards the stable steady state of minimum dissipation.

The Earth-atmosphere example effectively involves a 'series' system of energy conversions, and the state of maximum dissipation is also the state of maximum rate of energy transfer—a condition which has been the focus of several investigations in this and other branches of fluid mechanics^{5,6}.

GARTH W. PALTRIDGE

CSIRO Division of Atmospheric Physics,
Aspendale, Victoria,
Australia

Received 26 February; accepted 20 April 1979.

* Present address: CIRES, University of Colorado, Boulder, Colorado 80307.

1. Paltridge, G. W. *Q. Jl R. Meteor. Soc.* **104**, 927–954 (1978).
2. Lorentz, E. N. *The Nature and Theory of the General Circulation of the Atmosphere*, Ch. 5 (World Meteor. Organisation, Geneva, 1967).
3. Oort, A. H., *Mon. Weath. Rev.* **92**, 483–493 (1964).
4. Hunt, B. G. *J. Atmos. Sci.* **35**, 1133–1143 (1978).
5. Busse, F. H. *J. Fluid Mech.* **30**, 625–649 (1967).
6. Malkus, W. V. R. & Veronis, G. *J. Fluid Mech.* **4**, 225–260 (1958).
7. Haltiner, G. J. *Numerical Weather Prediction* (Wiley, New York, 1971).

Is there any scientific explanation of the paranormal?

THE apparent impossibility of the occurrence of 'paranormal' phenomena has not discouraged their extensive investigation, although there has not been any uniformly accepted validation or explanation by the scientific community. To clarify exactly how difficult ESP phenomena are to explain, it is necessary to place them in the framework of modern science. Explanations of the phenomena have been brought forward which have been claimed to make them more respectable. These explanations must also be looked at from the point of view of modern science and this paper is devoted to that task. In particular we wish to indicate that on theoretical grounds the only scientifically feasible explanation could be electromagnetism (EM) involving suitably strong EM fields. Thus we regard that this paper completes our earlier work¹ where we presented experimental results giving the level of the EM signals emitted by subjects when engaged in supposedly paranormal activity. These EM levels were many orders of magnitude lower than the ones we calculate here as needed to achieve paranormal effects. Taken together the two papers are a strong argument against the validity of the paranormal.

We will start our analysis by applying a criterion to which psychic phenomena and its explanations must conform, that of the conservation of energy (stemming as it does from the underlying symmetry of the forces of nature).

From energy conservation we may roughly assess the amount of energy transfer required to achieve a particular paranormal effect and hence assess its feasibility. The energies associated with reincarnation and precognition cannot be assessed at all; that for materialisation is expected to be of the order of megajoules if reasonably based on the energy required to break the molecular bonds in one mole of a normal material (500 kJ mol⁻¹). The energies needed for psychokinesis (PK), spoon-bending and poltergeists is of the order of joules, assuming that the phenomena occur by the gross effects to which they correspond, this being the movement or deformation of macroscopic objects. Finally the phenomena of dowsing, telepathy, distant-viewing, clairvoyance and faith-healing should need only of the orders of millijoules if it is assumed that they

require activation of certain portions of the brain of the subjects involved, though far more energy may need to be emitted by the source of these effects (as in telepathy or psychic healing). Only these latter phenomena can therefore be regarded as feasible, those requiring joules of energy being unlikely though faintly possible, whilst the first three phenomena contradict modern physics.

There are a range of natural energy sources which could be conjectured as being utilised to achieve so-called paranormal effects. Cosmic rays², nuclear beams³, neutrinos³, gravitons^{4,5} or other controls of gravity^{4,5} have all been suggested, though they all have such low power levels, survival value or high energy (by factors of about 10^{20}) required for control as to be completely irrelevant. A fifth force of nature^{6,7} explaining solely ESP phenomena would be purely *ad hoc*; faster-than-light particles⁸ and advanced potentials⁹ invoked to explain precognition have either severe theoretical difficulties¹⁰ or failed to appear¹¹⁻¹⁴. Higher dimensions for materialisation and precognition^{15,16} would be *ad hoc* as well as destroying the renormalisation program in quantum field theory.

None of the previous seven approaches suggest any mechanism by which human activity can bring about the desired paranormal phenomena. This is obviously not the case with theories in which the mind is independent of the body to some degree, such as those of Eccles^{17,18} and Popper^{19,20}, whose papers appear to contain no falsifiable statements.

Numerous attempts have been made to use the difficulties associated with measurements in quantum mechanics to explain paranormal phenomena²¹⁻²³. However, there seems to be considerable confusion in the motivation for these attempts, especially as the standard statistical interpretation of quantum mechanics^{24,25} has none of the claimed paradoxes. For paranormal effects to be explained in such a way they must contradict traditional quantum mechanics at some point, for which there is no other evidence.

From the viewpoint of modern science, the most reasonable approach to explain ESP, is electromagnetism (EM). We have previously reported¹ an extensive programme both of detection of EM radiation from subjects over nearly the whole spectrum and human body sensitivity to low-level EM radiation over a restricted, though very relevant, frequency range. We concluded that, for the subjects investigated and the phenomena associated with them (metal-bending, dowsing, psychic healing, telepathy and distant-viewing) EM cannot explain the effects. All the successful cases of PK¹ turned out to be completely explicable either by electrostatics or heat; we will not consider them here.

We shall now detail how difficult, if not impossible, it is for EM to account for ESP phenomena other than PK in the above-mentioned ways.

The human body is nearly a black body of temperature of 310 K that otherwise emits non-black-body EM radiation over a restricted range of physiological frequencies (d.c. to 5–6 kHz). As part of its black-body activity, the body radiates heat, mainly in the near infrared frequencies at wavelengths of $\sim 10\ \mu\text{m}$, the power level being $\sim 10\ \text{mW cm}^{-2}$. When investigating the problem of efficient emission of non-ionising EM radiation by the body in order to cause ESP effects of some kind, two crucial features have to be analysed: (1) spatial resolution, and (2) penetration depth into the body (or the emission from inside the body). Although the penetration depth inside the body increases as the frequency of the radiation decreases²⁶, if one allows the wavelength to increase beyond a certain limit the spatial resolution becomes very poor. This may not be crucial if it is just a question of emitting energy in all directions and so hopefully producing some paranormal effect in an arbitrary direction. It is critical if small objects are to be viewed by the emitted energy, such as in dowsing for a stream a metre or so wide, or clairvoyantly attempting to discern a picture a few centimetres wide on the face of a card.

We need only consider the attempt to use EM radiation for ESP purposes from the body or elsewhere at frequencies below 10 GHz. Above that frequency the penetration depth inside the

body is practically zero^{26,27} and EM emission will be determined completely by surface temperature. Small changes of the latter by several degrees have been reported to be possible to achieve by practice. These can produce at most only a 1% change of the black-body emission from the human body and so will be useless to explain any of the ESP phenomena requiring joules of energy mentioned earlier. If we think of dowsing, telepathy or distant-viewing as produced by an active or passive radar method using the infrared emission of the body, suitably pulsed (if that were possible), the reflected or received signal would be many orders of magnitude below the known sensitivity of the skin to infrared, such as in cases of dowsing and clairvoyance cited above and others typical in the ESP literature where positive results have been claimed. Temperature changes, both of healer and patient, do occur in the faith healing situation, but do not seem to be related to any healing process¹. Thus we conclude that skin emission or sensitivity cannot be relevant to ESP, and turn to frequencies below 10 GHz for which EM penetrates the human body appreciably.

For EM of radio frequencies lower than 100 MHz spatial resolution is poor due to a wavelength $\lambda > 3\ \text{m}$. This seems to rule out dowsing, distant-viewing and clairvoyance, due to inability to perceive objects smaller than 3 m across by such radiation. Dowers might still detect very small static magnetic fields, and underground objects can be associated with local magnetic field anomalies of the order of $10^{-5}\ \text{G}$. Our results¹ showed that dowers were unable to detect the presence of a magnet of $\sim 200\ \text{G}$. Therefore, it seems very unlikely that this mechanism is involved. At the lower end of the range, energy radiation is not a suitable concept, and static field ideas are more appropriate, especially in the non-black-body region below 5 kHz. Tests with healers¹ have indicated no changes in such levels to within millivolts of skin potential during healing sessions. Thus we can exclude this frequency range for faith healing also. Telepathy may be conjectured to occur by ELF/VLF^{28,29,30}, especially as there is the well known Schumann resonance at 8 Hz which allows for megametre propagation with little attenuation in the Earth-ionosphere cavity. However, the coupling of the human body to the long-wavelength radiation is especially crucial (a factor (human height/ λ)² $\sim 10^{-14}$ entering here). Kogan's analysis³⁰ does claim that somewhat higher frequencies are just feasible. However, values for the biocurrents induced in the receiver and required for the successful telepathic transmission of information are at most $10^{-12}\ \text{A}$ for a distance of a few metres between sender and receiver; for a 1-km separation this is reduced to $10^{-18}\ \text{A}$. The possible values of biocurrents produced by activity of a group of nerve cells is about $10^{-6}\ \text{A}$, so the induced currents are at least 10^6 (10^{12} for telepathy over 1 km) below that expected to noticeably bring a change of activity in a group of nerve cells. Even a single nerve cell, requiring at least $10^{-9}\ \text{A}$ to activate it, would have too low a current induced by such telepathic signals by a factor of 10^3 ; as at least 10^3 such cells would need to be activated to 'notice' the signal, the final factor of 10^6 is appropriate.

The more energetic paranormal phenomena, such as movements occurring in poltergeist activity involving heavy objects, are also readily ruled out, specifically due to the energy transfer needed. For metal bending, low-frequency fields above the breakdown strength of air would be required to produce bending or breaking in metal objects. For example, for a typical strip of metal with EM frequency of 10 kHz, power absorption by the metal of 1 mW would require an electric field of $10^6\ \text{V m}^{-1}$.

We finally turn to the more restricted frequency range of 100 MHz to 10 GHz. There is strong evidence²⁷ against non-thermal emission of radiation by the human body or even non-thermal effects³¹ induced in the body by very low levels of EM radiation. The power level emitted by the human body at microwave frequencies of 1–5 GHz is of the order of $10^{-14}\ \text{W}$. Therefore we can immediately discard the EM explanation for the more highly energetic phenomena that is materialisation, poltergeist, metal-bending and PK (translational motion of

objects). Clearly these are too low in energy by a factor of 10^{12} at least and up to 10^{20} for materialisation. One of the less energetic phenomena, psychic healing, would be also down by a factor of 10^{11} . It only remains to be seen how EM stands as regards the other less energetic phenomena, such as dowsing, telepathy, clairvoyance and distant-viewing. Two types of mechanisms can be envisaged for the less energetic phenomena: an active or a passive radar-like type.

Consider first site-dowsing. If we think of a dowser being able to detect underground water by means of an active radar-like mechanism, we see that radiation in the range 1–5 GHz at a power level of 10^{-14} W would be attenuated when penetrating through the soil, its skin depth being of a few cm. If a dowser used a passive radar-like mechanism, this would involve him sensing a cold spot (the underground stream) from an otherwise warmer environment at a depth, say, of 20 m. Our tests¹ on human sensitivity to low-level (1–5 mW) EM radiation in this frequency range indicate a lack of sensitivity by a factor of 10^{10} . The problems are even more difficult for map-dowsing because of the distance sometimes involved. The active radar mechanism can be ruled out, as the propagation distance of EM radiation of frequencies of 1–5 GHz is <100 km and dowsers have claimed map-dowsing successes up to a few thousand kilometres. The upper limit for round-the-world propagation of an EM signal is 70 MHz ($\lambda \geq 40$ m)³², and even then at power levels at least 10^{14} times greater than those from humans. There is also the problem of spatial resolution, as $\lambda \geq 40$ m and it could never account for the size of objects perceived (a few centimetres in some cases).

The cases of clairvoyance (direct or remote) and distant-viewing³³ also involve the problem of the size of the object perceived as compared with the wavelength that could be used by the subject ($\lambda \sim 6$ –300 cm). The minimum size $d_{\min}$ of an object observable at a distance R from an aerial of the size of the human body using radiation limited to a maximum frequency of 3 GHz for deep body penetration, satisfies $d_{\min} \geq R/40$, from standard diffraction theory. If $R \sim 1$ m, $d_{\min} \geq 2.5$ cm, contradicting the supposed ability in ref. 34 to read symbols down to 1/30 cm in size on a card a metre away. If $R \sim 4$ km, $d_{\min} \geq 100$ m, contradicting the distant-viewing abilities reported in the SRI tests³³ of observations of objects only 1 cm across.

Telepathy involves one subject (sender) transmitting encoded information to another. The receiver must then sense the incoming information, decode it and produce an output that can be compared with the original one sent to him. Here again, it is in the frequency range 1–5 GHz that this transmission would be most efficient. However, if we assume that the sender is using black-body EM radiation he is emitting in that frequency range, (necessarily above 10 kHz²⁷), the power level of the signal would be so low (by a factor of 10^5 for people 1 km apart) that it could never be sensed by the receiver. Our experiments have confirmed this by showing that human sensitivity is very poor at these frequencies¹.

We therefore conclude that neither EM nor any other scientific theory can explain any of the above mentioned ESP phenomena.

In particular there is no reason to support the common claim that there still may be some scientific explanation which has as yet been undiscovered. The successful reductionist approach of science rules out such a possibility except by utilisation of energies impossible to be available to the human body by a factor of billions.

We can only conclude that the existence of any of the psychic phenomena we have considered is very doubtful.

E.B. thanks the Sedwood Trust for financial support. We also thank Dr P. C. W. Davies for helpful discussions on advanced potentials.

J. G. TAYLOR
E. BALANOVSKI

Department of Mathematics,
King's College,
University of London,
London WC2, UK

Received 8 February; accepted 4 May 1979.

1. Taylor, J. G. & Balanovski, E. *Nature* **275**, 64 (1978).
2. Tromp, W. S. *Psychical Physics* (North Holland, Amsterdam, 1954).
3. Kervran, C. *1st European Parapsych. Conf.* Barcelona (1976); *Transmutations a faible energie*, (Librairie Malorine, Paris, 1972).
4. Dubrov, A. *Impact of Science on Society* **24**, 311 (1974).
5. Sarfatti, J. *J. Psych. Syst.* **1**, 3–10 (1974).
6. Wassermann, G. D. in *Extra-Sensory Perception* (Ciba Foundation Symposium, 1958).
7. Rolfe, W. G. *The Poltergeist* (New American Library Inc., New York, 1972).
8. Good, I. J. in *The Scientist Speculates: An Anthology of Parly-Baked Ideas* (eds Good, I. J., Mayne, A. J. & Maynard Smith J.) (London, Heinemann, 1962).
9. Puthoff, H. & Targ, R. *Proc. IEEE* **64**, 329–354 (1976).
10. Broido, M. M. & Taylor, J. G. *Phys. Rev.* **174**, 1606 (1968).
11. Davies, P. C. W. *J. Phys. A* **8**, 272 (1975).
12. Hoyle, F. & Narlikar, J. V. *Proc. R. Soc. A* **273**, 1 (1963).
13. Partidge, R. B. *Nature* **244**, 263 (1973).
14. Heron, M. L. & Pegg, D. T. *J. Phys. A* **7**, 1965 (1974).
15. Zollner, J. C. F. *Transcendental Physics* (Ballantyne, London, 1880).
16. Dunne, J. W. *An Experiment with Time* (Faber & Faber, London, 1927).
17. Eccles, J. C. *Facing Reality* (Springer, Berlin, 1970).
18. Eccles, J. *The Understanding Brain* (Academic, New York, 1973).
19. Popper, K. *Int. Conf. of Philosophy* **1**, Vienna 1968.
20. Popper, K. & Eccles, J. *The Self and its Brain* (Springer, Berlin 1979).
21. Walker, E. H. *Phys. Today* **24**, 39 (1971); *Consciousness and Quantum Theory in Psychic Exploration* (ed. Mitchell, E.) (Putnam's, London, 1974).
22. Mattuck, R. D. *Research on Parapsychology* (eds Morris et al. Scarecrow, Metuchen, 1977).
23. Mattuck, R. D. *J. Psych. Syst.* (in the press).
24. Ballentine, L. E. *Rev. Mod. Phys.* **42**, 358 (1970).
25. Belinfante, F. J. *Measurement and Time Reversal in Objective Quantum Theory* (Pergamon, Oxford, 1975).
26. Guy, A. W. & Johnson, C. C. *Proc. IEEE* **60**, 692 (1972).
27. Enander, B. & Larson, K. *Res. Rep. Trita. Tet.* 7505, 7506, 7701 (Royal Institute of Technology, Stockholm, 1976 & 1977).
28. Cazzamali, F. G. *Psych. Neurol., Estratto Fac.* **1** (I. Trimestre, 1933).
29. Vasiliev, L. L. *Experiments in Distant Influence* (ed. Gregory, A.) (Wedwood House, London, 1976).
30. Kogan, I. M. *Radio Energy* **21**, 75 (1966); **22**, 141 (1967); **23**, 122 (1968).
31. Presman, A. S. *Electromagnetic Fields and Life* (Plenum, New York, 1970).
32. Rumi, G. C. *Radio Sci.* **10**, 711–718 (1975).
33. Puthoff H. & Targ, R. *Nature* **252**, 606–607 (1974).
34. Rhine, J. B., Pratt, J. G., Smith, B. M., Stuart, C. E. & Greenwood, J. A. *ESP After 60 years* (Holt, New York, 1940).

A pachycephalosaurid dinosaur from Madagascar and a Laurasia–Gondwanaland connection in the Cretaceous

THE discovery of a pachycephalosaurid dinosaur in the Upper Cretaceous of Madagascar is reported here. This finding casts new light on the existence of a land connection (or several such connections) between Laurasia and Gondwanaland sometime in the Lower or Mid-Cretaceous. Hitherto this family of dome-headed ornithischians had only been known from the Northern Hemisphere: from the Lower Cretaceous of England¹ and the Upper Cretaceous of North America² and East Asia³.

Previously the evidence for a faunal connection between Laurasia and Gondwanaland during the Cretaceous was slight⁴. The frequently cited cosmopolitan distribution of the sauropod dinosaur *Titanosaurus* is based on very incomplete and often plainly non-diagnostic material and it is not improbable that several related genera are involved. The discovery of a dromaeosaurid in the Upper Cretaceous of Brazil has been reported⁴ but the lack of detailed information or illustrations makes evaluation of this record impossible. Hadrosaurine hadrosaurs are now known from several specimens from the Upper Cretaceous of Argentina^{5,6}; they previously represented the only, and more substantial, evidence for faunal exchange between Laurasia and Gondwanaland sometime in the Cretaceous. Fragmentary tyrannosaurid remains have been described from the Upper Cretaceous of India⁷ and South America^{8,9}. All other records of Laurasian dinosaur families from the Southern Hemisphere, which evolved solely within the Cretaceous period, are based on extremely fragmentary material which either lacks the diagnostic characters of the respective group or has been incorrectly determined⁹.

The supposed presence of a ratite bird, a member of a group otherwise restricted to the Southern Hemisphere, in the Upper Cretaceous of Mongolia is very doubtful; this form has been described on the basis of two badly crushed and incomplete

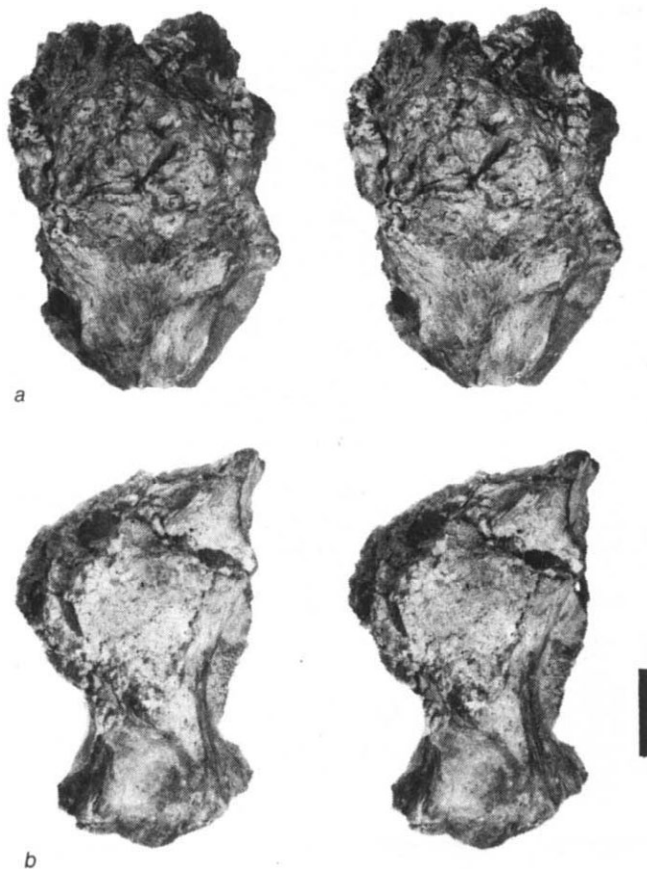


Fig. 1 *Majungatholus atopus*, new genus and species. MNHN.MAJ 4. Partial skull; stereo-photographs in dorsal (a) and lateral (b) view. Scale bar, 3 cm.

skulls¹⁰. It is quite unlike any known palaeognathous bird¹⁰ and in our opinion even its avian nature is not well established⁹.

The new find of a pachycephalosaurid comes from the Upper Cretaceous (probably Campanian) of the Majunga District in northwestern Madagascar. It was recently discovered by P.T. The specimen was collected at the beginning of this century; its true nature was not recognised at that time. The deposits in the Majunga region have yielded a rather diversified fauna of vertebrates^{9,11,12}, including a sauropod, a theropod and two teeth first referred to *Stegosaurus*¹³ but probably referable to some other group of ornithischians⁹.

Class: Reptilia
Order: Ornithischia
Suborder: Ornithopoda
Family: Pachycephalosauridae
Majungatholus, new genus.

Type species: *Majungatholus atopus*, new species.

Derivatio nominis: The generic name is derived from the name of the region where the specimen was found and the Latin *tholus*, 'dome'.

Diagnosis: A single dome-like thickening of the frontal region; frontal dome thick, with highly irregular, rugose dorsal surface. Parietal without significant thickening, participating in a well developed parieto-squamosal shelf. Very large supratemporal fenestrae. Olfactory portion of the braincase long; olfactory lobes ventrally enclosed by bone.

Majungatholus atopus, new species.

Holotype: Skull roof consisting of frontals and incomplete parietals and anterior part of the braincase (Fig. 1). Muséum National d'Histoire Naturelle, Paris, MNHN.MAJ 4 (formerly in the collections of the Ecole des Mines, Paris).

Locality and horizon: 'Grès de Maevarano'^{9,12}, Upper Cretaceous (probably Campanian), Majunga District, north-western Madagascar.

Derivatio nominis: The specific epithet is derived from the Greek *atopos*, 'strange' (or, in its original derivation, 'out of place'—a perhaps more appropriate meaning in the present context).

Diagnosis: As for the genus.

The specimen, a detailed anatomical description of which will be presented elsewhere⁹, evidently represents a dome-headed dinosaur referable to the Pachycephalosauridae. Like other pachycephalosaurids, it differs from the Ankylosauria in showing very long olfactory stalks and thickening of the cranial roof bones by upgrowth of the bones rather than by fusion of dermal ectopic elements¹⁴. Furthermore, in ankylosaurs the upper temporal fenestrae are always closed¹⁵.

Majungatholus represents a distinctive type of the dome-headed ornithischians; its single frontal dome and its very large upper temporal fenestrae are unique features. The earliest known pachycephalosaurid, *Yaverlandia bitholus*, from the Lower Cretaceous (Wealden) of the Isle of Wight¹, shows some resemblance in having large supratemporal fenestrae and thickening restricted to the frontal region; it differs, however, in having a small dome on each frontal and in being much smaller. The only Upper Cretaceous pachycephalosaurid species showing a certain similarity is a flat-headed form referred to *Stegoceras* from the Campanian of Alberta, Canada¹. In this species the supratemporal fenestrae are large and a parieto-squamosal shelf is well developed; the frontals, however, are but slightly thickened and the species is smaller than *Majungatholus atopus*. A peculiar feature only present in *Majungatholus* and the two Laurasian species mentioned above is a median circular depression on the postero-dorsal surface of the frontals.

Majungatholus can be derived from a *Yaverlandia*-like form and probably represents a lineage evolving in geographical separation from Northern Hemisphere pachycephalosaurids.

The presence of a pachycephalosaurid in Madagascar and of other 'Laurasian' dinosaur families in the Southern Hemisphere during the Late Cretaceous can be explained in two different ways. First, dispersal from Laurasia could be assumed as all these families have their greatest diversity and most primitive species in the Northern Hemisphere. The Tethyan sea belt separated the two super-continent¹⁶ during the Cretaceous but it was at least in part a fairly shallow epicontinental sea and passage routes may have become available on several occasions. The exact locations of some of these connections may never be known as many potential areas have subsequently been affected by Cretaceous and Tertiary orogeneses. Cox⁴ has suggested the existence of a filter route in the Central American or Caribbean region to account for the presence of a hadrosaur in Argentina. But it can also be argued that the existence of similar dinosaurs in both hemispheres is the result of vicariance events (the breakup of Pangea). The evidence presently available does not permit a choice between the two hypotheses. It is clear that faunal exchange by land connections between Laurasia and Gondwanaland was possible at least during the first half of the Cretaceous, possibly even at a later date.

The photographs for Fig. 1 were prepared by A. H. Coleman of Harvard University.

HANS-DIETER SUES

Museum of Comparative Zoology and
Department of Biology,
Harvard University,
Cambridge, Massachusetts 02138

PHILIPPE TAQUET

Institut de Paléontologie,
Muséum National d'Histoire Naturelle,
8 Rue de Buffon, 75005 Paris, France

Received 27 March; accepted 18 April 1979.

1. Galton, P. M. *J. Paleont.* **45**, 40–47 (1971).
2. Brown, B. & Schlaikjer, E. M. *Bull. Am. Mus. nat. Hist.* **82**, 115–150 (1943).
3. Maryńska, T. & Osmólska, H. *Palaeont. pol.* **33**, 45–102 (1974).
4. Cox, C. B. *J. Biogeogr.* **1**, 75–94 (1974).
5. Casamiquela, R. M. *Ameghiniana* **3**, 285–312 (1964).
6. Brett-Surman, M. K. *Nature* **277**, 560–562 (1979).
7. Chatterjee, S. J. *Paleont.* **52**, 570–580 (1978).
8. Woodward, A. S. *Proc. Zool. Soc. Lond.* 169–184 (1901).
9. Sues, H.-D. *J. Paleont.* (in the press).
10. Elzanowski, A. *Nature* **264**, 51–53 (1976).
11. Depéret, C. *Bull. Soc. géol. Fr.* (3) **21**, 176–194 (1896).
12. Russell, D., Russell, D., Taquet, P. & Thomas, H. C. *r. somm. Soc. géol. Fr.* **5**, 204–208 (1976).
13. Piveteau, J. *Bull. Soc. géol. Fr.* (4) **26**, 33–38 (1926).
14. Sues, H.-D. *N. Jb. Geol. Paläont. Mh.* 459–472 (1978).
15. Coombs, W. P. Jr *Palaeontology* **21**, 143–170 (1978).
16. Hallam, A. (ed.) *Atlas of Palaeobiogeography* (Elsevier, Amsterdam, 1973).

Growth of host root establishes contact with parasitic angiosperm *Boschniakia hookeri*

MEMBERS of the family Orobanchaceae are root parasites attached underground to suitable hosts. In some cases attachment is effected by special mechanisms of the parasite. For example, the mature seed of *Orobanche*, stimulated chemotropically by diffusion of compounds from the host, germinates and its extending radicle establishes contact with the host^{1,2}. We report here a previously unknown mechanism in *Boschniakia hookeri*, a perennial parasite found only on the roots of certain members of the Ericaceae in a narrow westerly distribution on the North American continent. We found that the connection of the parasite to the host is due ultimately to the behaviour of the host root, which grows through the testa of the parasite's seed into contact with the embryo.

We studied a population of *B. hookeri* growing in Kitsap County, Washington. Field observations showed that dispersal of the seeds depends in part on the properties of the seeds themselves, and in part on environmental factors. The mature seeds, as they are exposed in the ripe capsule, are relatively large (1.8 ± 0.2 mm), non-wettable, and of low specific gravity ($0.2-0.25$). These properties result from the structure of the testa, a light, dry, honeycomb-like sheath surrounding the small embryo (about 0.4×0.3 mm). Thus the seeds are easily displaced and chance movements (rain drops, animal contact) release them

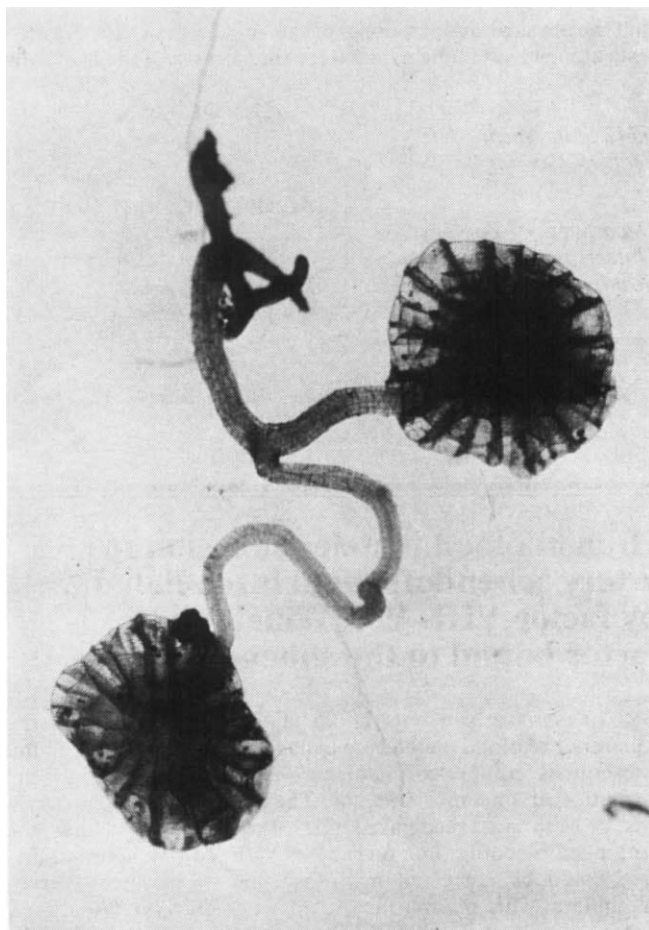


Fig. 2 Branched roots of *Gaultheria shallon* with the root-tips growing through the outer testa cells towards the embryo. Material identical with seeds in Fig. 1.

from the capsule. Further dispersal is probably mainly through natural pores in the soil, or through tunnels made by small mammals (such as moles and mice). Occasional underground hoarding of seeds by mammals may also be a factor, although of secondary importance. Apparently the site of germination and the resulting tuber is determined merely by where seed and host root meet. Some tubers were found just beneath the surface, while in other cases tubers were as much as 25 cm below the surface—a depth of 10–15 cm was most common.

The embryo does not contain specialised organs³ and apparently does not take an active part in the contact with the root. Connection is established by the young root-tip of the host, growing into and through the large outer cells of the testa (Figs 1, 2) until it presses against the embryo of the parasite. Later, the tissues of the parasitic embryo and the host merge. Cell division of the seed results in the formation of a tuber, in which vascular and meristematic tissues differentiate. This process can occur within 2–4 months from the time when the root and seed first come into contact. Apparently germination is not strictly dependent on the season, for young tubers are found in a continuous series of sizes, with a very close correlation between their dimensions and the diameter of the host roots.

The attraction between host root and parasite seems highly effective, although unexplained. For example, in germination experiments up to 75% of seeds placed on the roots of young *Arctostaphylos uva-ursi* were later found with the host roots growing into them. It is possible that the host roots are attracted by exudates from the embryo or testa. But as host root-tips can penetrate non-living organic particles (such as fragments of tree bark) and inorganic particles in the soil (such as artificially introduced Perlite, a heat-expanded volcanic glass), it may be

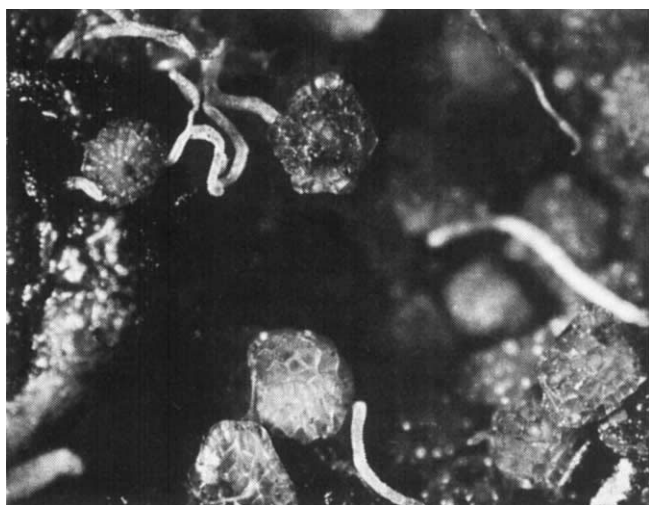


Fig. 1 Duff taken from below *Gaultheria shallon*. Many seeds of *Boschniakia hookeri* are present, as well as young roots of *Gaultheria*. The characteristic outer layer of deeply pitted cells of the testa (seed coat) makes it easy to recognise the *Boschniakia* seeds.

that the physical characteristics of the outer cells of the parasite's testa alone result in the contact-seeking growth of the host root.

SIGURD OLSEN

6542 55th Avenue NE,
Seattle, Washington 98115

INGRITH DEYRUP OLSEN

Department of Zoology,
University of Washington,
Seattle, Washington 98195

Received 12 February; accepted 18 April 1979.

1. Vaucher, J. P. *Mém. Mus. Hist. nat. Paris* **10**, 261–273 (1823).
2. Kuijt, J. *The Biology of Parasitic Flowering Plants* (University of California Press, Berkeley and Los Angeles, 1969).
3. Tiagi, B. *Bot. Notiser* **116**, 81–93 (1963).

Human blood platelet adhesion to artery subendothelium is mediated by factor VIII–Von Willebrand factor bound to the subendothelium

ONE of the early events in the haemostatic process is the adherence of blood platelets to collagen of the subendothelium. Subsequent platelet-to-platelet adhesion then leads to the formation of a haemostatic plug. The role of a plasma factor in this process was recognised after the discovery^{1,2} that the prolonged bleeding time of patients with Von Willebrand disease could be corrected by transfusion of plasma or cryoprecipitate. This plasma factor, then named Von Willebrand factor, was later identified as factor VIII (anti-haemophilic

factor A (ref. 3)). The current concept is that Von Willebrand factor and factor VIII are present in a complex, which can be dissociated in certain conditions. Despite this progress, little is yet known about the role of the factor VIII–Von Willebrand factor complex (factor VIII–VWF) in the haemostatic plug formation. Tschopp *et al.* showed that the Von Willebrand factor was required for normal adhesion of platelets to rabbit subendothelium in a perfusion chamber using blood from patients with Von Willebrand disease⁴. The role of factor VIII–VWF in platelet adhesion is also important for the development of atherosclerosis. In this process, adherence of platelets is followed by degranulation and concomitant secretion of factors causing intimal migration and proliferation of smooth muscle cells^{5,6}. Defective adherence of platelets was shown to protect against experimental atherosclerosis^{7,8}. We present here evidence that factor VIII–VWF binds to subendothelium of human arteries and that subendothelium-bound factor VIII–VWF mediates the platelet adhesion.

A perfusion chamber as developed by Baumgartner⁹ was used to study the effect of factor VIII–VWF on platelet adhesion to subendothelium, but using components of human origin only. Blood from healthy donors was anticoagulated with citrate. Platelets were washed in Krebs–Ringer solution, labelled¹⁰ with ⁵¹CrO₄ and incubated with aspirin to prevent platelet–platelet interaction¹¹ (for details see Fig. 1). Factor VIII–VWF was isolated from human cryoprecipitate¹² and labelled with ¹²⁵I by the lactoperoxidase method¹³. Subendothelium of human postmortem renal arteries after the first bifurcation was used; arteries with macroscopically observable atherosclerotic lesions were discarded. Perfusions were carried out at a shear rate comparable to that in arteries¹⁴.

In the first set of experiments the specificity of the effect of factor VIII–VWF on the adhesion of ⁵¹Cr-platelets was studied. After perfusion for 5 min at 37 °C with normal plasma containing red cells (haematocrit 40%) and ⁵¹Cr-platelets (2.5 × 10⁵ per



Fig. 1 Light microscopical picture (×1,100) of human subendothelium after perfusion with aspirin-treated platelets. No platelet interaction could be detected, only a monolayer of platelets covering the subendothelium. All arteries used had subendothelial fibrosis. Fixation, embedding, cutting and staining was after Baumgartner¹⁵. The arteries were obtained from autopsies usually 12 h after death. They were exposed to air for a few minutes to remove endothelial cells completely, and the lumen was wiped with a bud probe. Finally, the lumen was rinsed with 0.2 M Tris buffer (pH 7.4). On light microscopical examination no endothelial cells could be detected. The reconstituted blood used for the perfusions was anticoagulated with 19 mmol citrate (plasma concentration). Platelet-rich plasma (PRP) was obtained from whole blood by centrifugation (10 min at 190g, 20 °C). To one volume of PRP one volume of a Krebs–Ringer solution (4 mM KCl, 107 mM NaCl, 20 mM NaHCO₃ and 2 mM Na₂SO₄) containing 19 mM citrate and 0.5% glucose (pH 5.0) was added which gave a final pH of 6.1. Platelets were spun down (10 min, 500g, 20 °C) and resuspended in 2 ml Krebs–Ringer solution (with citrate and glucose, pH 6.1) four times. The second time 1 μCi CrO₄ per ml was added to the resuspended platelets (about 10⁶ per μl) and centrifugation was carried out after 20 min at 20 °C. The third time, 10 μM aspirin was added for 15 min at 37 °C. After the fourth washing over 97% of the ⁵¹CrO₄ was intracellular. Platelet poor plasma (PPP) was obtained from PRP by centrifugation (15 min, 3,000g, 4 °C). Plasma contamination of the red cells was avoided by three washings in saline containing 2% glucose. The labelled platelets, PPP and the red cells were reconstituted to give 2.5 × 10⁵ platelets per μl plasma and a red cell concentration of approximately 40%. The pH was adjusted to 7.3. The perfusates were incubated for 5 min at 37 °C before the perfusions were carried out. A peristaltic tube pump (Verder) was used for the perfusions. The average flow rate was 135 ml min⁻¹ and the average shear rate was calculated to 805 s⁻¹. All perfusions were carried out for 5 min at 37 °C. Immediately after perfusion the system was rinsed with 40 ml 0.2 M Tris buffer (pH 7.4). A segment of approximately 2.5 mm was cut from both ends of the artery. It was calculated that a 2.5-mm broad area of turbulence would occur where the flow met the edge of the vessel wall¹⁶. The remaining part (~0.5 cm²) was used for registration of ⁵¹Cr (Trigamma 600, Baird Atomic).

Table 1 Platelet adherence to subendothelium in media with and without factor VIII-VWF

	Platelet no. ($\times 10^5$) per cm^2 subendothelium (mean $\pm$ s.d.)	
	Without addition of factor VIII-VWF	Addition of 1 U factor VIII-VWF per ml
Normal plasma	—	71.5 $\pm$ 11.0 (7)
Haemophilia A plasma	—	67.2 $\pm$ 5.8 (5)
Von Willebrand plasma*	28.8 $\pm$ 8.2 (11)	73.3 $\pm$ 8.7 (6)
Plasma after cryoprecipitation†	24.5 $\pm$ 7.6 (6)	70.2 $\pm$ 5.3 (5)
Human albumin solution (4.0%)	26.2 $\pm$ 8.7 (7)	77.5 $\pm$ 8.0 (5)
Dextran 40.000 solution (3.3%)	31.2 $\pm$ 3.1 (5)	60.3 $\pm$ 7.5 (5)

Number of perfusions in parentheses.

* Factor VIII coagulant activity 14%, ristocetin cofactor assay 5% and factor VIII-related antigen 7%.

† Factor VIII coagulant activity 15%, ristocetin cofactor assay 1% and factor VIII-related antigen 15%.

μl plasma), 70×10^5 platelets per cm^2 subendothelium were detected. When normal plasma was substituted by plasma from patients with Von Willebrand disease or by supernatant plasma after cryoprecipitation, the platelet adhesion was markedly decreased. A similar decrease was found on substitution of normal plasma by solutions of human albumin or dextran, which indicated that no proteins other than factor VIII-VWF play a part in the platelet adhesion. Indeed, addition of factor VIII-VWF to these plasma substitutes corrected the platelet adhesion to the same level as obtained with normal plasma (Table 1). This normalisation was reached in a relatively small range (0.5–0.8 U per ml of factor VIII-VWF as measured in the ristocetin cofactor assay¹⁷).

The mechanism by which factor VIII-VWF promotes the adhesion of platelets to the subendothelium was studied in the following way. First, perfusions (5 min, 37 °C) were carried out using solutions of human albumin and factor VIII-VWF (1 U ml^{-1} including 10% ^{125}I -factor VIII-VWF). These experiments indicated a binding of 10^{-3} U factor VIII-VWF per cm^2 subendothelium. Then, double-perfusion experiments were carried out to study the influence of bound factor VIII-VWF on the platelet adhesion. Subendothelium was perfused with human albumin solutions containing various amounts of ^{125}I -factor VIII-VWF, followed by a second perfusion with red cells and ^{51}Cr -platelets in albumin solutions without factor VIII-VWF. Figure 2 shows the relationship between the concentration of factor VIII-VWF and the binding of this protein to the subendothelium (A) and the dependence of the platelet adhesion on the amount of factor VIII-VWF bound (B). A nearly normal adhesion was obtained after binding of about 10^{-3} U factor VIII-VWF per cm^2 subendothelium (4×10^9 molecules per cm^2). Removal of factor VIII-VWF from the subendothelium during the second perfusion could not be detected. It should be mentioned that even complete desorption would yield a negligible concentration of factor VIII-VWF in the second perfusate.

From these experiments, carried out with human components in conditions which mimic the *in vivo* situation after vessel wall injury, we conclude (1) that for normal adhesion factor VIII-VWF is the only plasma protein necessary, and (2) that factor VIII-VWF bound to subendothelium is a highly effective inducer of platelet adhesion.

The results strongly suggest that binding of factor VIII-VWF to subendothelium is an important step in the first phase of the haemostatic plug formation. This is supported by ultrastructural studies of haemostatic plugs in Von Willebrand patients^{18,19} showing a disturbed adherence of platelets to the vessel lips and the presence of free platelet aggregates in the wounds. The effect of factor VIII-VWF on platelet adhesion may also be important for the development of atherosclerosis, as was concluded by Fuster *et al.*⁸, who showed that pigs with severe Von Willebrand disease are much less susceptible to atherosclerosis than control pigs.

Furthermore, the present results are interesting in view of the synthesis of factor VIII-VWF-related antigen (with Von

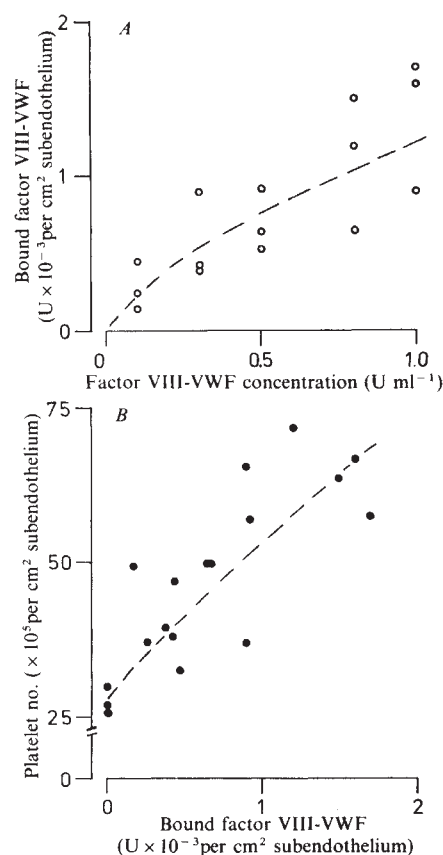


Fig. 2 Factor VIII-VWF accumulation and the subsequent platelet adhesion on human subendothelium in a double-perfusion experiment. A, Factor VIII-VWF bound to the subendothelium in the first perfusion (○) with different factor VIII-VWF concentrations. B, Number of platelets adhered to the subendothelium in the second perfusion (●), with the corresponding amount of pre-bound factor VIII-VWF. Dashed curves are power curve fits according to the formula $y = ax^b + c$, where in A: $a = 1.23$, $b = 0.70$, $c = 0.0$ and the correlation coefficient $r = 0.85$, and in B: $a = 26.0$, $b = 0.86$, $c = 26.7$ and $r = 0.84$. In the first perfusion a Krebs-Ringer solution containing 19 mM citrate, 2.4 mM CaCl_2 , 0.5% glucose and 4% human albumin (recrystallised, Sigma A 9511) with different concentrations of purified factor VIII-VWF was perfused for 5 min, at 37 °C. 10% of the purified factor VIII-VWF was labelled with ^{125}I . After the first perfusion, the system was rinsed with 135 ml 0.2 M Tris buffer (pH 7.4). The second perfusion (5 min, 37 °C) was then started immediately. This perfusate contained the Krebs-Ringer solution as used in the first perfusion but without factor VIII-VWF. Platelet and red cell concentrations were similar to the concentrations, as indicated in Fig. 1 legend. After the second perfusion, the system was rinsed with 40 ml 0.2 M Tris buffer (pH 7.4). The middle part of the artery was cut from the rod and used for registration of ^{51}Cr and ^{125}I .

Willebrand activity but without coagulant activity) in the endothelial cells²⁰. On vessel wall injury, this antigen may be bound to the subendothelium and subsequently mediate the adhesion of platelets. A local haemostatic effect of factor VIII-VWF-related antigen present in the vessel wall has been suggested by Bloom *et al.*²¹ on the basis of immunohistochemical studies of the intima and by Holmberg *et al.*²² who reported the absence of this antigen in the vessel wall of severely affected Von Willebrand patients.

We thank Dr H. R. Baumgartner (Roche) for his advice concerning the perfusion procedure and Dr L. van Gorp and H. Westendorp (St Antonius Hospital, Utrecht) for the supply of renal arteries. Financial support was obtained from the Netherlands Heart Foundation (grant no. 76090).

K. S. SAKARIASSEN
P. A. BOLHUIS
J. J. SIXMA

Department of Haematology,
University Hospital,
Utrecht, The Netherlands

Received 7 November 1978; accepted 17 April 1979.

- Nilsson, I. M., Blombäck, M. & Blombäck, B. *Acta med. scand.* **164**, 263–278 (1959).
- Cornu, P., Larrieu, M. J., Caen, J. & Bernard, J. *Br. J. Haemat.* **9**, 189–202 (1963).
- Bouma, B. N., Wiegierick, Y., Sixma, J. J., Van Mourik, J. A. & Mochtar, I. A. *Nature new Biol.* **236**, 104–106 (1972).
- Tschopp, T. B., Weiss, H. J. & Baumgartner, H. R. *J. Lab. clin. Med.* **83**, 296–300 (1974).
- Stemerman, M. B. & Ross, R. *J. exp. Med.* **136**, 769–789 (1972).
- Ross, R., Glomset, J., Kariya, B. & Harker, L. A. *Proc. natn. Acad. Sci. U.S.A.* **71**, 1207–1210 (1974).
- Harker, L. A., Ross, R., Slichter, S. J. & Scott, C. R. *J. clin. Invest.* **58**, 731–741 (1976).
- Fuster, V. *et al. J. clin. Invest.* **61**, 722–730 (1973).
- Baumgartner, H. R. & Haudenschild, C. *Ann. N.Y. Acad. Sci.* **201**, 22–36 (1972).
- Abrahamsen, A. F. *Scand. J. Haemat.* **5**, 53–63 (1968).
- Evans, C., Packham, M. A. & Nishizawa, E. E. *J. exp. Med.* **128**, 877–894 (1968).
- Van Mourik, J. A. & Mochtar, I. A. *Biochim. biophys. Acta* **221**, 677–679 (1970).
- Over, J. *et al. J. clin. Invest.* **62**, 223–234 (1978).
- Wells, R. & Goldstone, J. in *Rheology of Biological Systems* (eds Gabelnick, H. L. & Litt, M.) 42 (Thomas, Springfield, 1973).
- Baumgartner, H. R. *Microvascular Res.* **5**, 167–179 (1973).
- Lighthill, M. J. *J. Fluid Mech.* **52**, 475–479 (1972).
- Howard, M. A. & Firkin, B. G. *Thromb. Diath. Haemorrhage* **26**, 362–369 (1971).
- Jørgensen, L. & Borchgrevink, C. F. *Acta path. microbiol. scand.* **60**, 55–82 (1964).
- Hovig, T. & Stormorken, H. *Acta path. microbiol. scand. Suppl.* **248**, 105–122 (1974).
- Jaffe, E. A., Hoyer, L. W. & Nachman, R. L. *Proc. natn. Acad. Sci. U.S.A.* **71**, 1906–1909 (1974).
- Bloom, A. L., Giddings, J. C. & Wilks, J. C. *Nature new Biol.* **241**, 217–219 (1973).
- Holmberg, L., Mannucci, P. M., Turesson, I., Ruggeri, Z. M. & Nilsson, I. M. *Scand. J. Haemat.* **13**, 33–38 (1974).

Calcium conductance of acetylcholine-induced endplate channels

SEVERAL lines of evidence indicate that an influx of Ca^{2+} ions accompanies transmitter activation of the postsynaptic membrane at the neuromuscular junction^{1–5}. We report here that we have now investigated the elementary characteristics of the Ca^{2+} current by bathing muscles in CaCl_2 solutions containing no Na^+ ions, so that Ca^{2+} is the only ion available to carry any appreciable inward current^{1,3,6,7}. The single channel current (i) and mean lifetime (τ) of the postsynaptic channels were then determined by the technique of noise analysis^{8,9}. It was found that τ in the Ca^{2+} solution was shorter than in normal Ringer, but that the voltage dependence of τ was unchanged. The elementary current showed a non-linear voltage dependence, unlike the linear dependence in normal solution⁸.

Experiments were carried out on the frog (*Rana temporaria*) sartorius muscle, at a temperature of 5–6 °C. The muscle was strongly stretched to reduce contraction and its associated artefacts. The normal bathing solution contained (in mM): NaCl, 120; KCl, 2; CaCl_2 , 1.8; HEPES, 4 (pH 7.2). Calcium solutions contained: CaCl_2 , 82 or 160; KCl, 2; HEPES, 4 (pH 7.2): care was taken to avoid contamination with Na^+ . A standard two-point voltage clamp was used at the endplates, and a third external micropipette filled with 2 M acetylcholine (ACh) was used to apply ACh ionophoretically to the endplate. A fast

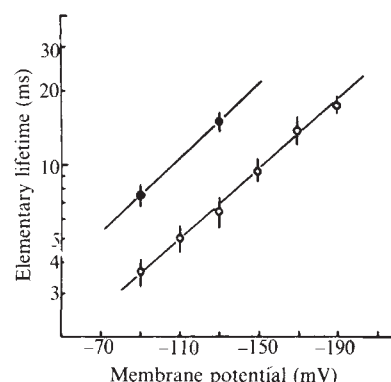


Fig. 1 Membrane potential dependence of the ACh-induced single channel lifetime, τ , measured from muscles bathed in normal Ringer solution (●) and in isotonic or hypertonic Ca^{2+} solutions (○). No differences were seen in τ between isotonic and hypertonic Ca^{2+} solutions, and the results shown are pooled data from both. Lines were fitted by eye, and error bars give ± 1 s.d. Five endplates were examined in Ringer solution, and seven in CaCl_2 solutions. Values were obtained during steady ionophoretic application of ACh to voltage-clamped endplates, by fitting the membrane current fluctuation spectra to a lorentzian curve. Background spectra were subtracted before fitting. τ was calculated from the half power frequency f_c of the spectra as $\tau = 1/(2\pi f_c)$. About 15 noise segments of 512 points were obtained at each potential, at a digitisation rate of 500 or 1,000 Hz, and were used to compute an average spectrum.

Fourier transform method was used to analyse the frequency components of the current fluctuations during steady ACh application, and i and τ were determined as previously described^{8,9}.

Records were usually obtained from several fibres in normal Ringer solution, and then from the same, and additional fibres, after changing to Ca^{2+} solution. Resting membrane potentials were higher in isotonic and hypertonic Ca^{2+} solutions (mean values: normal Ringer 85.1 ± 5.8 mV, Ca^{2+} solution 103 ± 7.2 mV: all deviations are ± 1 s.d. unless otherwise stated), and the input resistance of the fibres was increased (normal Ringer ~ 350 k Ω ; Ca^{2+} solution 1.32 ± 0.38 M Ω). Pipettes with large tip diameters were used to reduce background noise, and the values given above may be artificially low because of fibre damage during penetration. The frequency of miniature endplate potentials was elevated after changing to Ca^{2+} solution, but fell to nearly zero after 2–3 h, when ACh noise recordings could be made. Miniature endplate currents (m.e.p.cs) were reduced in duration and size in Ca^{2+} solution (see also ref. 3), mean amplitudes at -130 mV being 2.86 ± 0.39 nA in normal Ringer, and 0.97 ± 0.37 nA in Ca^{2+} solution.

In normal Ringer, ionophoretic application of large doses of ACh to the endplate caused local contractions, and these became much larger in isotonic Ca^{2+} solution. For example, a barely visible contraction, involving two sarcomeres, was seen with a synaptic charge flux of about 1 nC. Because of this difficulty in avoiding movement artefacts, some experiments were carried out in hypertonic Ca^{2+} solution (160 mM CaCl_2), which helped to reduce contraction. To guard against the possibility that part of the synaptic current was carried by an influx of Cl^- ions⁷, recordings were made from two fibres in a bathing solution of 82 mM Ca-cyclamate. M.e.p.cs were recorded in this solution, with the same amplitudes as in 82 mM CaCl_2 solution, and there was little difference in reversal potential. Also K^+ ion influx made no appreciable contribution to the synaptic current, as no significant changes in equilibrium potential or single channel currents were seen in fibres bathed in a solution containing only 160 mM CaCl_2 .

A variety of divalent cations in addition to Ca^{2+} are able to carry a synaptic current. We have recorded m.e.p.cs and ACh-induced currents from fibres bathed in isotonic solutions of MgCl_2 , MnCl_2 , SrCl_2 and CoCl_2 (all solutions also contained 2 mM KCl and 4 mM HEPES). This contrasts with the

behaviour in several other systems, where Mg^{2+} , Mn^{2+} and Co^{2+} are known to block Ca^{2+} fluxes¹⁰⁻¹³. The drug D600 has also been shown to block Ca^{2+} fluxes in a variety of preparations¹¹⁻¹³. We found that at a concentration of $200 \mu g ml^{-1}$ D600 decreased the frequency of m.e.p.s in isotonic Ca^{2+} , and did not abolish the ACh sensitivity of the muscle membrane. However, with ionophoretic ACh application only brief transient responses could be recorded, and a period of a few minutes was required for recovery of the response to a second pulse. This effect of D600, which was also present in normal Ringer, may be attributable to a blocking of open channels, or to an increase in desensitisation.

Figure 1 shows the voltage dependence of mean lifetime (τ) of ACh-induced channels in Ca^{2+} solution, plotted on semi-logarithmic coordinates. The data are fitted well by a straight line, indicating that τ increases exponentially with membrane hyperpolarisation. The voltage constant (potential shift to give an e -fold change in τ) was 62 mV. τ increases exponentially with hyperpolarisation in normal Ringer⁸, and the Ca^{2+} data closely parallel a line drawn through control measurements in normal Ringer, although shifted in a hyperpolarised direction by about 46 mV. The shortening of τ in Ca^{2+} solution is surprising, as an increase in Ca^{2+} concentration from 1 to 10 mM has been reported to slow the decay phase of m.e.p.s¹⁴. However, we have been unable to confirm these results, and found that this increase in Ca^{2+} concentration produced in some experiments a shortening of m.e.p.c. decay, and in others no detectable change.

Figure 2 shows the voltage dependence of the elementary current, derived from noise analysis, in hypertonic Ca^{2+} solution. The relationship is not linear throughout, indicating that the elementary conductance (γ) varies with membrane potential. Closely similar behaviour was seen in experiments using isotonic Ca^{2+} solution. This is unlike the situation in normal Ringer, where γ is independent of potential⁸, and in our experiments had a value of 29.3 ± 2.7 pS (s.e.m.). In hypertonic Ca^{2+} solution the voltage dependence of i was approximately linear at potentials more hyperpolarised than -110 mV, and the slope of this segment gave a value for γ of 6.25 pS. Noise measurements at potentials between equilibrium and -100 mV were very difficult to obtain, but an average value for γ of ~ 2.5 pS was estimated for this segment.

Our results confirm previous observations¹⁻⁵ showing that activation of the postsynaptic membrane with ACh causes an increase in permeability to Ca^{2+} ions, and further indicate that Mg^{2+} , Mn^{2+} , Sr^{2+} and Co^{2+} ions are also able to cross the channels induced by ACh. However, the conductances for these ions are small, and at -90 mV, for example, the elementary conductances in both isotonic $MgCl_2$ and $CaCl_2$ solutions are about 10 times less than in normal Ringer.

Divalent ions might cross the endplate membrane (1) nonspecifically through the 'ordinary' channels normally used

by Na^+ and K^+ , or (2) through a separate Ca^{2+} -specific channel. The apparently identical voltage dependence of i for Ca^{2+} and Na^+ currents, together with the failure of Mg^{2+} , Mn^{2+} , Co^{2+} and D600 to block the m.e.p.s, favour (1), but do not entirely exclude (2). The reason for the differences in conductance properties in Na- or Ca-Ringer is not clear, and the non-linear dependence of the elementary Ca^{2+} current on membrane potential is not predicted by the Goldman-Hodgkin-Katz model of ion permeation¹⁵⁻¹⁸. However, differences also exist between various divalent ions themselves, and we find that, unlike Ca^{2+} , the elementary Mg^{2+} current has a linear voltage dependence over the range examined (-50 to -130 mV).

The decrease in channel lifetime in Ca^{2+} solutions is in the opposite direction to that expected from the effect of Ca^{2+} ions in screening negative surface charges on the muscle membrane¹⁴, and suggests that Ca^{2+} ions have an additional effect on the channel. It may be that in addition to direct effects of various ions on the receptors and their environment, the lifetime of transmitter-induced channels is influenced by the nature of the ions which permeate them¹⁹.

We thank Sir Bernard Katz for helpful discussion, and Abbot laboratories for D600.

P. D. BREGESTOVSKI*
R. MILEDI
I. PARKER

Biophysics Department,
University College London,
Gower Street, London WC1, UK

Received 19 February; accepted 8 May 1979.

* Present address: Institute of Biological Physics, Puschino, Moscow region, USSR.

1. Takeuchi, N. *J. Physiol., Lond.* **167**, 141-155 (1963).
2. Jenkinson, D. H. & Nicholls, J. G. *J. Physiol., Lond.* **159**, 111-127 (1961).
3. Katz, B. & Miledi, R. *J. Physiol., Lond.* **203**, 459-487 (1969).
4. Evans, R. H. *J. Physiol., Lond.* **240**, 517-533 (1974).
5. Miledi, R., Parker, I. & Schalow, G. *J. Physiol., Lond.* **269**, 11-13P (1977).
6. Heuser, J., Katz, B. & Miledi, R. *Proc. R. Soc. B178*, 407-415 (1971).
7. Takeuchi, N. *J. Physiol., Lond.* **167**, 128-140 (1963).
8. Anderson, C. R. & Stevens, C. F. *J. Physiol., Lond.* **235**, 655-691 (1973).
9. Katz, B. & Miledi, R. *J. Physiol., Lond.* **224**, 665-699 (1972).
10. Katz, B. & Miledi, R. *J. Physiol., Lond.* **203**, 459-487 (1969).
11. Sanchez, J. A. & Stefani, E. *J. Physiol., Lond.* **283**, 197-209 (1978).
12. Hagiwara, S. & Nakajima, S. *J. gen. Physiol.* **49**, 793-806 (1966).
13. Baker, P. F., Meves, H. & Ridgeway, E. B. *J. Physiol., Lond.* **231**, 511-526 (1973).
14. Cohen, I. & Van der Kloot, W. *Nature* **271**, 77-79 (1978).
15. Goldman, D. E. *J. gen. Physiol.* **27**, 37-60 (1943).
16. Hodgkin, A. L. & Katz, B. *J. Physiol., Lond.* **108**, 37-77 (1949).
17. Meves, H. *Pflügers Arch. ges. Physiol.* **304**, 215-241 (1968).
18. Lewis, C. A. *J. Physiol., Lond.* **286**, 417-445 (1979).
19. Van Heiden, D., Hammill, O. P. & Gage, P. W. *Nature* **269**, 711-713 (1977).

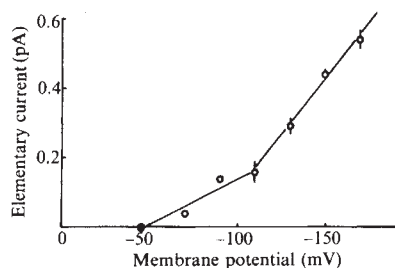


Fig. 2 Membrane potential dependence of the single channel current i , measured from fibres bathed in 160 mM $CaCl_2$ solution. Lines were fitted by eye. Points shown with error bars are means from five fibres, and bars are ± 1 s.e.m. Points at -90 and -70 mV are measurements from single fibres. The equilibrium potential ($\bullet$) was measured from the reversal potential of the ACh response in seven fibres, and was -44.4 mV (s.d. ± 1.1 mV). Elementary currents were calculated from the total spectrum variance, after subtraction of background variance. The alternative technique of computing i from the zero frequency asymptote of the spectra gave values in good agreement.

Anti-HLA-A,B,C monoclonal antibodies with no alloantigenic specificity in humans define polymorphisms in other primate species

THE study of evolutionary relationships by immunological comparison of homologous proteins from different species has been limited by the complexity and variability of conventional heteroantisera¹. These problems can be avoided by using selected panels of monoclonal antibodies, specific for single antigenic determinants, which would enable detailed immunochemical descriptions of homologous proteins to be made². The human HLA-A,B,C antigens and their homologues provide a particularly interesting system for this type of analysis³. Although they are extraordinarily polymorphic within species, many features of their structure, including association with β_2 -microglobulin, have been highly conserved during evolution⁴⁻⁶. A preliminary survey showed that monoclonal antibodies with specificity for the HLA-A,B,C- β_2 -microglobulin complex reacted with primate species but not with non-primates⁷. We have therefore

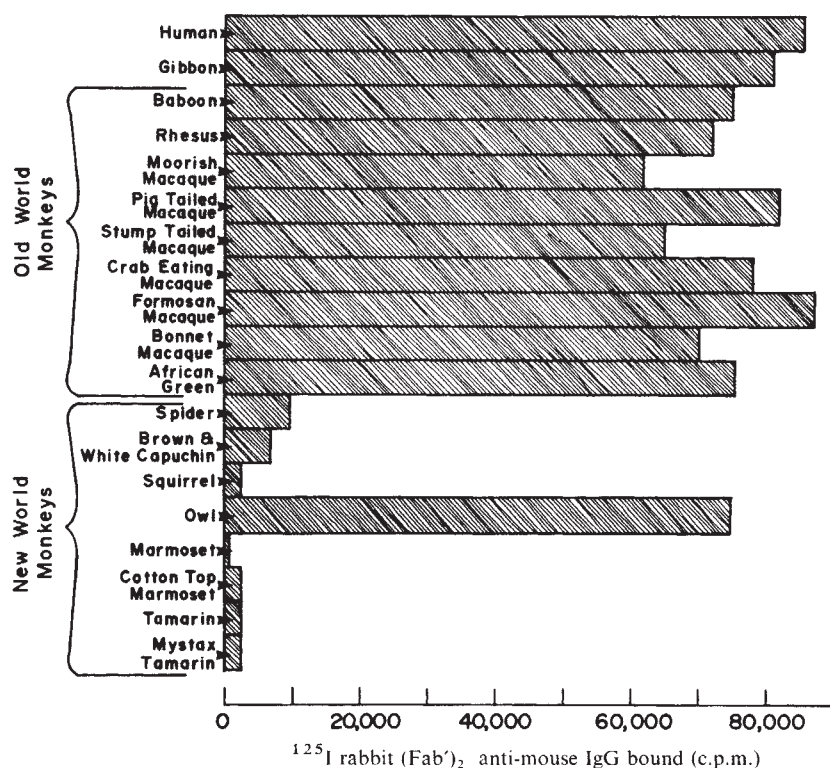


Fig. 1 Indirect trace binding of W6/32 antibody to lymphocytes of different primate species. The assay was basically as described in ref. 14. Peripheral blood lymphocytes (5×10^5) in a volume of 50 μ l were incubated with 50 μ l of a 1/100 dilution of ascites fluid from mice carrying the W6/32 hybrid cell as a tumour, for 1 h at 4°C. After three washes the cells were incubated with 2×10^6 c.p.m. of 125 I-labelled rabbit (Fab')₂ anti-mouse IgG for 1 h at 4°C, washed four times and counted for radioactivity.

concentrated on comparing primate species with a panel of anti-A,B,C monoclonal antibodies. We report here the finding that two different anti-HLA-A,B,C monoclonal antibodies with no demonstrable polymorphic specificity in humans define polymorphisms in owl and spider monkeys, thus demonstrating the potential of these reagents for fine structure antigenic mapping.

Results obtained from a limited number of species indicated that the anti-HLA-A,B,C monoclonal antibody W6/32 reacted with apes and Old World monkeys whereas the anti- β_2 -microglobulin monoclonal antibody BBM1 only reacted with the higher apes, chimpanzee and gorilla^{7,8}. The indirect binding of W6/32 to peripheral lymphocytes from individual animals of 19 primate species is shown in Fig. 1. Positive reactions were found with all species of Old World monkeys and the only ape (gibbon) tested. All New World species, with the exception of the owl monkey, were negative. The BBM1 antibody, as expected (ref. 7, and data not shown), was negative against all non-human species tested.

To confirm the presence of the W6/32 antigenic determinant on owl monkey lymphocytes four animals were individually tested. Three gave negative reactions and the single animal tested in the original screen was still positive. Subsequent typing of 40 owl monkeys showed 11 positive and 29 negative. When lymphocytes from W6/32 positive and negative animals were assessed for their capacity to bind W6/32 antibody (Fig. 2), positive animals gave titration curves indistinguishable from human and rhesus lymphocytes, in contrast to negative animals, which gave no significant binding at any cell number and were indistinguishable from spider monkey lymphocytes used as the negative control.

The owl monkey is widely distributed throughout South America, and regional variations in coat colour and chromosomal morphology have been described⁹. The polymorphic reaction pattern of W6/32 antibody was found to correlate exactly with chromosomal differences. Nine animals with karyotype VI from Bolivia and two with karyotype VII from Peru were all positive, six animals with karyotype I from Brazil and 23 animals with karyotypes II, III, IV and V from Columbia were all negative (Fig. 3 and Table 1).

Other anti-HLA-A,B,C monoclonal antibodies which showed no polymorphic reactions in humans¹⁰ were screened

against owl monkey lymphocytes of different karyotype. Three different reaction patterns were seen: (1) negative on all cells, (2) identical to W6/32 antibody, and (3) a polymorphic reaction pattern different from the W6/32 antibody. Three antibodies of the last class were identified which all gave similar reaction patterns. Positive reactions were seen with all animals of karyotypes I, II, III, IV, V and VII whereas polymorphic reactions were found with animals of karyotype VI (Table 1). One antibody (PA2.5) was selected for further study and its specificity on all human cells confirmed. PA2.5 antibody reacted equally well with a panel of 20 HLA-A, B, C-bearing human B-cell lines and this binding was equally inhibitable by purified preparations of both HLA-A locus (A2 and A28) and HLA-B locus (B7 and B40) antigens¹¹.

As the initial screen (Fig. 1) with the W6/32 antibody against different species involved single animals and by chance, an owl monkey of karyotype VII had been selected, it became important to re-evaluate the reactions on other species. Groups of three New World species (squirrel, spider, cebus) and a single Old World species (rhesus) were tested against W6/32 and PA2.5 antibodies and the results are shown in Table 1. The original conclusions concerning the W6/32 antigenic determinant were confirmed. It was absent on all cebus, squirrel and spider monkeys and present on all rhesus monkeys tested. The

Table 1 Indirect binding reactions of W6/32 and PA2.5 antibodies on primate lymphocytes

Species	No. of animals tested	W6/32		PA2.5	
		Positive	Negative	Positive	Negative
Spider monkey	12	0	12	8	4
Owl monkey (karyotype VI)	9	9	0	3	6
Owl monkey (karyotype I, II, III, IV, V)	29	0	29	29	0
Owl monkey (karyotype VII)	2	2	0	2	0
Squirrel monkey	20	0	20	20	0
Cebus (brown and white capuchin) monkey	15	0	15	15	0
Rhesus monkey	14	14	0	14	0

PA2.5 antibody gave positive reactions with all rhesus, squirrel and cebus monkeys but was polymorphic in the spider monkey.

The polymorphisms defined by the two monoclonal antibodies in the owl monkey were quite distinct. W6/32 antibody reactions correlated exactly with known differences in karyotype, coat colour and geographical origin. This result is similar to that found for wild mice where individual demes of geographically separated populations had characteristic H-2 antigens¹². The PA2.5 antibody defined a polymorphism within a presumably interbreeding group of animals that were identical for karyotype, coat colour and W6/32 antibody reactivities. Similarly, the polymorphism defined by the PA2.5 antibody in the spider monkey was not correlated with other known parameters.

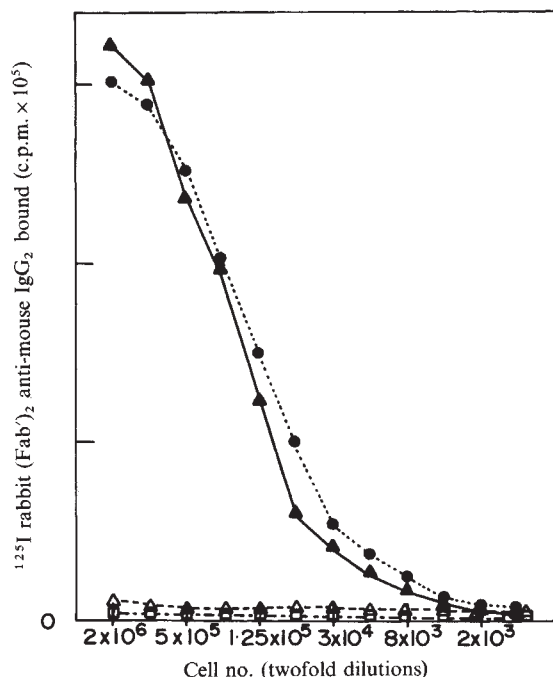


Fig. 2 Indirect trace binding of W6/32 antibody to human, spider monkey and owl monkey lymphocytes as a function of cell number. Assays were as described in the Fig. 1 legend except different cell numbers were used. ●, Owl monkey with karyotype VI; □, owl monkey with karyotype I; ▲, human used as a positive control; △, spider monkey used as negative control.

It is not unexpected that non-polymorphic anti-HLA-A,B,C monoclonal antibodies define polymorphism in other primate species. Highly conserved regions of HLA-A,B,C glycoproteins will not be immunogenic in mice, and it is likely that monoclonal antibodies made in mice will recognise parts of the molecule that can tolerate amino acid substitutions without loss of function and thus have the potential for polymorphism within a species.

Most non-polymorphic anti-HLA,B,C monoclonal antibodies we have produced bind to cells of non-human primate species. Comparison of their reactions with different species has provided a simple way of defining different antigenic determinants of the HLA-A,B,C- β_2 -microglobulin complex. The W6/32 antibody defines a determinant generally restricted to apes and Old World monkeys, the PA2.5 antibody defines a determinant present on apes, Old World monkeys and most New World monkeys. These results indicate that both antigenic determinants originated before the divergence of New and Old World monkeys and it will be interesting to see if they can be detected in tree shrews and the more primitive species of primates.

The high frequency with which antibodies of similar specificity are produced suggests that the number of antigenic determinants is small. This is probably a reflection of the $\approx 70\%$

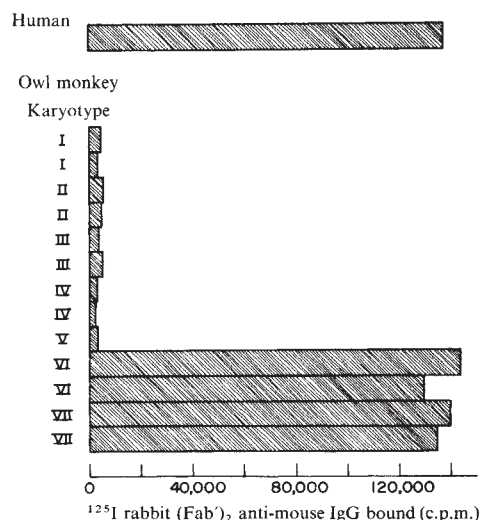


Fig. 3 Indirect trace binding of W6/32 antibody to lymphocytes from owl monkey with seven different karyotypes. Assays were as described in the Fig. 1 legend.

amino acid sequence homology between HLA-A,B,C glycoproteins and the H-2D,K glycoproteins of the mice used to make the antibodies^{5,6}. The finding that HLA-A,B,C antigens and their homologues in higher primates are antigenically similar in comparison to mouse indicates a very high degree of amino acid sequence homology. The equivalence with which many anti-HLA monoclonal antibodies bind to cells of non-human primate species and the use of insolubilised W6/32 antibody in the purification of HLA-A,B,C antigens¹³ suggest that the homologous antigens of other species may be purified by this method and made available for structural characterisation.

We thank Jack Thomson and Joanne McLean for technical assistance and Dr Walter Bodmer for helpful discussions. F.M.B. was supported by a Marshall Scholarship from the Marshall Aid Commemoration Commission. P.P. was supported by a Junior Fellowship from the Society of Fellows, Harvard University and thanks Dr Jack L. Strominger for provision of laboratory facilities. This research was supported by grants from the UK MRC, and the US NSF and NIH.

PETER PARHAM*

Sidney Farber Cancer Institute,
Harvard Medical School,
44 Binney St, Boston, Massachusetts 02115

P. K. SEHGAL

New England Regional Primate Center,
Harvard Medical School,
1 Pine Hill Drive, Southborough, Massachusetts 01772

FRANCES M. BRODSKY

Genetics Laboratory,
University of Oxford,
South Parks Road, Oxford, UK

Received 19 March; accepted 24 April 1979.

* To whom correspondence should be addressed.

- Williams, C. A. in *Evolutionary and Genetic Biology of Primates* (ed. Buettner-Janusch, J.) 25 (Academic, New York, 1964).
- Köhler, G. & Milstein, C. *Nature* **256**, 495 (1975).
- Götze, D. (ed.) *The Major Histocompatibility System in Man and Animals* (Springer, Berlin, 1977).
- Dick, H. M. in *Histocompatibility Testing*, 157 (Munksgaard, Copenhagen, 1978).
- Coligan, J. E. *et al. Molec. Immun.* **16**, 3 (1979).
- Ploegh, H. L. *et al. in Biological Markers of Neoplasia* (ed. Ruddon, H.) 201 (Elsevier, Amsterdam, 1978).
- Brodsky, F. M., Bodmer, W. F. & Parham, P. *Eur. J. Immun.* (in the press).
- Barnstable, C. J. *et al. Cell* **14**, 809 (1978).
- Ma, N. S. F. *et al. Lab. Anim. Sci.* **26**, 1022 (1976).
- Parham, P. & Bodmer, W. F. *Nature* **276**, 397 (1978).
- Parham, P. *et al. J. biol. Chem.* **252**, 7555 (1977).
- Klein, J. *Transplantation* **13**, 291 (1972).
- Parham, P., Barnstable, C. J. & Bodmer, W. F. *J. Immun.* (in the press).
- Williams, A. F. *Contemp. Topics molec. Immun.* **6**, 83 (1977).

Functional separation *in vivo* of both antigens encoded by H-2 subregion and non-H-2 loci

ANTIGENS coded for by the H-2K, H-2D and H-2I region of the major histocompatibility complex (MHC) of the mouse initiate different *in vitro* responses. I region-coded antigens activate mainly Lyt 1⁺ T cells to produce a proliferative response in mixed lymphocyte cultures (MLC), whereas K and D region-coded antigens predominantly stimulate Lyt 2⁺ T cells to become cytolytic effector cells¹. Proliferative responses in MLC and generation of cytolytic T cells *in vitro* can also be induced by minor histocompatibility antigens. For example, Mls-locus products can induce positive MLC responses, but not the generation of cytolytic T cells², whereas H-Y antigen does give rise to anti-H-Y cytotoxic T-cell responses *in vitro*, following *in vivo* priming³. Functional *in vivo* studies have shown that H-2K, H-2D and H-2I differences can account for graft rejection⁴ and for mortality in graft-versus-host (GvH) reactions⁵. Differences in only minor histocompatibility antigens are sufficient to cause graft rejection⁶, and lethal GvH reaction after allogeneic bone marrow transplantation⁷. Similarly, immunisation with only H-2 or only non-H-2 antigens can induce a state of delayed-type hypersensitivity (DTH) to the immunising antigen, which can be measured with the footpad swelling test⁸. So far, such *in vivo* experiments have shown little or no discrimination between

responses to H-2 subregion antigens and responses to non-H-2 antigens. We have used a DTH assay to study the occurrence of T-effector cells after *in vivo* immunisation with different histocompatibility antigens. We show here that DTH T-effector cells generated in GvH and host-versus-graft (HvG) reactions are specific for largely different sets of histocompatibility antigens, with selective stimulation by H-2I and Mls-locus antigens in GvH conditions.

We have developed a simple DTH assay which can be used to measure the development of T-effector cells induced by histocompatibility antigens during a GvH reaction^{9,10}. The assay is based on secondary transfer of lymphoid cells from animals undergoing a GvH reaction, and subsequent testing of the secondary recipients for DTH reactivity to the histocompatibility antigens which evoked the GvH reaction. Briefly, lymphoid cells from mice which have been irradiated and reconstituted with allogeneic spleen cells, are transferred intravenously (i.v.), at various intervals after reconstitution, into normal secondary recipients syngeneic to the original spleen cell donor. The secondary recipients are challenged in the right hind foot with 2×10^7 spleen cells syngeneic to the original irradiated recipients. The DTH response to this challenge is measured as the difference in thickness of the hind feet 24 h later. The specific increase in foot thickness is calculated as the relative increase in foot thickness of the immune mice minus the relative increase in foot thickness of control mice which have only received the challenge. The swelling of challenged control mice varies between 12 and 20%. Using this assay we studied the capacity of

Table 1 Anti-host immune reactivity in lethally irradiated recipients after spleen cell transplantation across H-2-subregion differences

Inoculum	Recipient	H-2 subregion coding for immunising antigen*	Challenge	H-2 subregion coded antigens of cells used for challenge	DTH Response
B10.A	B10.AQR	K	B10.AQR	K	3.2 ± 1.6
B10.AQR	B10.A × B10.T(6R)	K + I	B10.A × B10.T(6R)	K + I	35.5 ± 1.7
B10.AQR	B10.A × B10.T(6R)	K + I	B10.T(6R)	I	32.2 ± 2.9
B10.AQR	B10.A × B10.T(6R)	K + I	B10.A	K	3.4 ± 3.0
A.TH	A.TL	I(+Tla)	A.TL	I(+Tla)	31.4 ± 2.6
A.TH	A.SW	D(+Tla)	A.SW	D(+Tla)	-0.3 ± 1.4

GvH reactions were elicited by i.v. injection of 5×10^7 spleen cells (inoculum) into lethally irradiated recipient mice. Recipients were lethally irradiated (750 rad) within 4 h before reconstitution with the spleen cell inoculum; 5 d after reconstitution two-thirds of the total cell yield obtained from spleen, inguinal, axillary and mesenteric lymph nodes of a recipient mouse were transferred i.v. to a secondary recipient which was syngeneic to the original spleen cell donor. Lymphoid cells from all primary recipients of a particular combination were pooled before secondary transfer. Challenge was carried out with 2×10^7 (50 µl) spleen cells (treated with mitomycin C (Kyowa Hakko Kogyo) $100 \mu\text{g ml}^{-1}$ for 30 min at 37 °C) injected s.c. into the instep of the right hind foot of the secondary recipients. At 24 h after challenge DTH responses were measured. Control mice only received a challenge dose. The swelling of these control mice was 12–20%. DTH responses are expressed as the specific percentage increase in foot thickness, as described previously⁹. Figures represent the arithmetic mean ± 1 s.e.m. of five mice.

* H-2-subregion difference between spleen cell donor and irradiated recipient mice. The origin of H-2 subregions (K, I-A, I-B, I-C, D) for the strains are as follows: A.TH, ssssd; A.SW, sssss; A.TL, skkdd; B10.A, kkkdd; B10.AQR, qkkdd; B10.T(6R), qqqqd.

Table 2 Anti-host immune reactivity in lethally irradiated recipients after spleen cell transplantation across Mls-locus differences

Inoculum	Recipient	Peak MLC response	Challenge	Mls locus-coded antigen of cells used for challenge	Peak DTH response
BALB/c	BALB/c × DBA/2	24,000 ± 850 (4)	DBA/2	Mls ^a	25.5 ± 0.7 (5)
BALB/c	BALB/c × DBA/2	—	B10.D2	Mls ^b	0.5 ± 1.2 (5)
DBA/2	BALB/c × DBA/2	630 ± 280 (3)	BALB/c	Mls ^b	10.6 ± 2.2 (3)
AKR/FuRdA	C3H/f	22,300 ± 1300 (3)	C3H/f	Mls ^c	25.8 ± 1.8 (7)
AKR/FuRdA	C3H/f	—	B10.BR	Mls ^b	2.9 ± 1.1 (7)
C3H/f	AKR/FuRdA	38,000 ± 2900 (4)	AKR/FuRdA	Mls ^a	26.7 ± 2.0 (7)
C3H/f	AKR/FuRdA	—	B10.BR	Mls ^b	1.9 ± 1.7 (7)

GvH reactions were elicited by i.v. injection of 5×10^7 spleen cells (inoculum) into lethally irradiated recipient mice. The designation for H-2 haplotype and Mls locus for the strains are as follows: BALB/c, H-2^d, Mls^b; DBA/2, H-2^d, Mls^a; BALB/c × DBA/2, H-2^{d/d}, Mls^{b/a}; B10.D2, H-2^d, Mls^b; AKR/FuRdA, H-2^k, Mls^a; C3H/f, H-2^k, Mls^c; B10.BR, H-2^k, Mls^b. Recipients were lethally irradiated (850 rad) before reconstitution. At different intervals after reconstitution a number of spleen cells equivalent to one spleen was transferred i.v. into secondary recipients which were syngeneic to the original spleen cell donors. Peak one-way MLC responses are expressed as specific counts per min. Figures represent the arithmetic mean ± 1 s.e.m. of a quadruplicate microculture. The day of peak MLC response is given in parentheses. Each culture consisted of 5×10^5 responder spleen cells and 5×10^5 stimulator (treated with 25 µg mitomycin C per ml) spleen cells. Activity of responder cells alone was 250–1,200 c.p.m. and activity of mitomycin C-treated stimulator cells 50–300 c.p.m. Background activity was 35 c.p.m. Details of the technique will be published elsewhere¹⁵. Challenge was carried out as described in Table 1 legend. Peak DTH responses are expressed as the specific percentage increase in foot thickness. Figures represent the arithmetic mean ± 1 s.e.m. of five mice. The day of peak DTH response is given in parentheses.

Table 3 Anti-graft immune reactivity after immunisation with spleen cells differing at H-2 subregions or minor histocompatibility loci

Immunisation	Responder	Challenge	H-2 subregion or Mls locus-coded antigens of cells used for challenge	DTH Response
B10.AQR	B10.A	B10.AQR	K	39.2 ± 2.8
B10.AQR	B10.T(6R)	B10.AQR	I	38.0 ± 3.2
A.TL	A.TH	A.TL	I(+Tla)	33.6 ± 3.0
A.SW	A.TH	A.SW	D(+Tla)	26.6 ± 2.3
BALB/c × DBA/2	BALB/c	DBA/2	Mls ^a	39.2 ± 2.2
BALB/c × DBA/2	BALB/c	B10.D2	Mls ^b	28.4 ± 3.0
BALB/c × DBA/2	DBA/2	BALB/c	Mls ^b	44.4 ± 2.8
AKR/FuRdA	C3H/f	AKR/FuRdA	Mls ^a	31.0 ± 2.2
AKR/FuRdA	C3H/f	B10.BR	Mls ^b	22.1 ± 2.0
C3H/f	AKR/FuRdA	C3H/f	Mls ^c	34.6 ± 2.4
C3H/f	AKR/FuRdA	B10.BR	Mls ^b	28.8 ± 1.5

Immunisation was carried out with 1×10^7 spleen cells s.c., equally distributed over the inguinal areas by means of a 28-gauge needle. For the origin of H-2 subregions and Mls locus see legends to Tables 1 and 2. Challenge was carried out as described in the Table 1 legend. DTH responses were assayed 5 d after s.c. immunisation of the responder mice and expressed as the specific percentage increase in foot thickness. Figures represent the arithmetic mean $\pm$ 1 s.e.m. of five mice.

H-2 subregion-coded antigens and minor histocompatibility antigens to induce such T-effector cells after (semi-)allogeneic spleen cell transplantation in lethally irradiated mice. These results were correlated with the capacity of the same antigens to induce DTH-related T-effector cells during a HvG reaction.

In lethally irradiated congenic mice, reconstituted with allogeneic spleen cells which were different at the whole or part of the H-2 complex, it could be demonstrated that the transferable anti-host DTH reactivity is directed exclusively to the I region of the H-2 complex; K and D subregion-coded antigens do not induce anti-host DTH T-effector cells (Table 1). H-2 compatible, non-H-2-incompatible spleen cell transplantation in lethally irradiated hosts revealed that antigens encoded by the Mls locus can also elicit the development of a transferable anti-host-directed DTH reactivity (Table 2). Mls^a and Mls^c coded antigens initiated both a positive MLC response and a distinct GvH-related DTH reactivity. Mls^b-locus antigens, on the other hand, were not able to initiate *in vitro* proliferation; this was associated with a marginal and short-lasting GvH-related DTH reactivity (Table 2).

Minor histocompatibility antigens other than Mls-locus products did not contribute to the expression of the anti-host DTH reactivity. This was shown by B10.D2 challenge of secondary BALB/c recipients of cells activated to DBA/2 non-H-2 antigens, and by B10.BR challenge of C3H/f and AKR recipients of cells activated to AKR and C3H/f, respectively. In this experiment the cells used for challenge had different Mls-locus antigens from the irradiated recipients. As there is extensive cross-reactivity in various mouse strains between non-H-2 alloantigens other than Mls-locus products¹¹, challenge with H-2-compatible, Mls locus-different spleen cells would demonstrate the contribution of these non-H-2 alloantigens in the expression of GvH-related DTH reactivity. No significant anti-host DTH reactivity could be elicited after challenge with B10.D2 or B10.BR spleen cells (Table 2). Therefore, non-H-2 alloantigens other than Mls locus-coded products probably do not evoke anti-host DTH reactivity. Apparently, Mls-locus products and I-region antigens have a similar capacity for inducing anti-host-directed DTH T-effector cells. K and D region antigens and non-H-2 antigens other than Mls-locus products are much less able to do so.

In contrast, in HvG reactions induced by subcutaneous (s.c.) injection of allogeneic spleen cells, K, D and I region-coded antigens as well as non-H-2 histocompatibility antigens could give rise to anti-graft DTH T-effector cells (Table 3). Also, minor histocompatibility antigens other than Mls-locus products contribute to the expression of the anti-graft DTH reactivity (Table 3). According to the work of Vadas *et al.*¹² there may be two subsets of T cells mediating DTH: Lyt 1⁺ and Lyt 2⁺ T cells, which are restricted by the I or the K and D regions of the MHC, respectively. Probably, in GvH conditions only Lyt 1⁺ DTH

T-effector cells are activated (anti-H-2I and anti-Mls response) and in HvG conditions both Lyt 1⁺ and Lyt 2⁺ DTH T-effector cells are activated. This might be related to the frequency of T cells reactive to the various histocompatibility antigens^{13,14}.

We thank Drs O. Vos, Th.H. van der Kwast, H. von Boehmer and K. Fischer-Lindahl, and Mr N. H. C. Brons for valuable contributions. The investigation was supported by the Netherlands Foundation for Medical Research (FUNGO).

EMILE A. J. WOLTERS

Department of Cell Biology and Genetics,
Erasmus University, PO Box 1738 Rotterdam,
The Netherlands

ROBERT BENNER

Basel Institute for Immunology,
487 Grenzacherstrasse,
Basel, Switzerland

Received 29 January; accepted 2 May 1979.

1. Bach, F. H., Bach, M. L. & Sondel, P. M. *Nature* **259**, 273-281 (1976).
2. Festenstein, H. in *Immunobiology of Bone Marrow Transplantation* (eds Dupont, B. & Good, R. A.) 13-21 (Grune and Stratton, New York, 1976).
3. Simpson, E. & Gordon, R. D. *Immun. Rev.* **35**, 59-75 (1977).
4. Klein, J. in *Biology of the Mouse Histocompatibility-2 Complex*, 151-177 (Springer, Berlin, 1975).
5. Klein, J. & Chiang, C. L. *J. Immun.* **117**, 736-740 (1976).
6. Graff, R. J. & Bailey, D. W. *Transplant. Rev.* **15**, 26-49 (1973).
7. Korngold, R. & Sprent, J. *J. exp. Med.* **148**, 1687-1698 (1978).
8. Kon, N. D. & Klein, P. A. *J. Immun.* **117**, 413-415 (1976).
9. Wolters, E. A. J. & Benner, R. *Transplantation* **26**, 40-45 (1978).
10. Wolters, E. A. J. & Benner, R. *Transplantation* **27**, 39-42 (1979).
11. Bevan, M. J. *J. exp. Med.* **142**, 1349-1364 (1975).
12. Vadas, M. A., Miller, J. F. A. P., Whitelaw, A. M. & Gamble, J. R. *Immunogenetics* **4**, 137-153 (1977).
13. Ford, W. L., Simmonds, S. J. & Atkins, R. C. *J. exp. Med.* **141**, 681-696 (1975).
14. Lindahl, K. F. & Wilson, D. B. *J. exp. Med.* **145**, 508-522 (1977).
15. Wolters, E. A. J., Brons, N. H. C., Van der Kwast, T. H. & Benner, R. (in preparation).

A candidate for the permeability pathway of the outer mitochondrial membrane

UNLIKE most other membranes, the outer mitochondrial membrane seems to be freely permeable to various small molecules and has therefore been called 'leaky'. The intactness of this membrane is demonstrated by observations¹⁻⁴ which show it to be impermeable to large polymers. This 'sieving' property is consistent with the presence of size-selecting units such as channels. Parsons and co-workers^{5,6} have published electron micrographs showing what looked like channels in the outer membrane of mitochondria (particularly those from plants), and recent X-ray diffraction studies by Mannella and Bonner⁷ have confirmed the existence of structures of similar dimensions to those seen in the electron micrographs. I now report the extraction of channel-forming material from the outer membrane of rat liver mitochondria. This material has properties which are

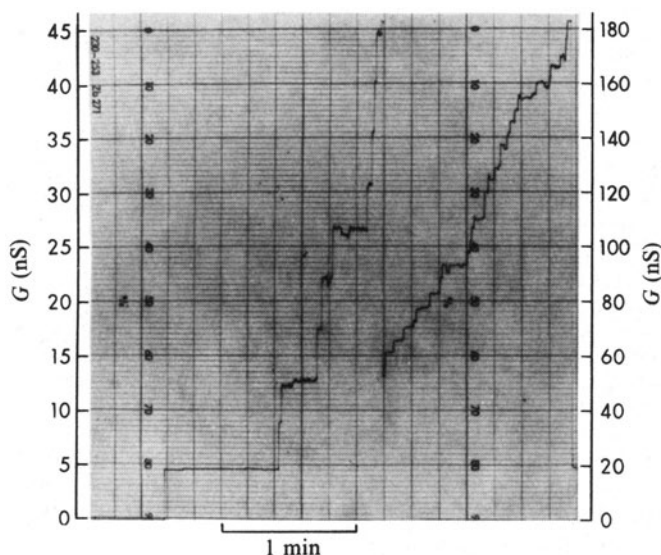


Fig. 1 Incorporation of Triton X-100-solubilised VDAC into a planar bilayer. An unmodified planar lipid bilayer was made (see Table 1 legend) with a conductance of ~ 10 pS (membrane was 0.15 mm in diameter) in the presence of 1 M KCl, 5 mM CaCl_2 . Five microlitres of a 2% Triton solution containing VDAC from rat liver mitochondria (a purified fraction—the method of purification will be described elsewhere) was added to the front compartment. A transmembrane voltage of 10 mV (the rear compartment being positive) was applied continuously using the voltage clamp set-up previously described⁸. After a few minutes, the insertion of channels is observed as distinct increases in membrane conductance. The channels are of uniform size. In the middle of the record the scale was reduced by a factor of 4, so the scale on the right-hand side of the graph refers to the right-hand side of the record while that on the left-hand side refers to the left side of the record. The chart speed was constant throughout.

essentially identical with those reported previously⁸ for voltage-dependent anion-selective channels (VDAC) from mitochondria of *Paramecium*. VDAC could be responsible for the permeability of the outer membrane to small molecules and may be involved in regulating mitochondrial metabolism.

The VDAC channels previously described (from *Paramecium* mitochondria) displayed the following properties after insertion into planar lipid bilayers: (1) The permeability produced in the bilayer is due to channels of uniform size. (2) These channels conduct anions better than cations of comparable size and charge. (3) The channels switch to lower conducting states (close) when the transmembrane voltage is increased both in the positive and negative direction: the voltage range at which this switching from high to low conductance occurs is narrow so that the voltage dependence is steep. These properties characterise VDAC and distinguish it from all other conductances that have been observed in bilayers. These properties are constant from preparation to preparation and, remarkably, from species to species. Hence, the term VDAC is used here to refer to all the channels displaying these characteristic properties.

Channel-forming activity was assayed in two ways as described in Table 1 legend and conductance in the bilayer was checked for the characteristic properties described above. In the first assay method the sample (suspended in hexane) was layered on the surface of the aqueous phases to generate monolayers from which the bilayer was formed (no detergent was used). The numbers of VDAC channels in the resulting bilayer are a measure of VDAC activity. Results obtained with this method are shown in the third column of Table 1. The second method involved extracting a sample with Triton X-100 and adding an aliquot of this extract to the medium bathing an unmodified lipid bilayer. The rate of VDAC channel insertion from the medium per unit time is a measure of the VDAC content of the original sample (last column in Table 1). Both methods yield VDAC channels with the same characteristic properties. Therefore the presence of the detergent has no discernible effect on VDAC.

Table 1 shows that mitochondria from a wide variety of eukaryotes have VDAC channels. To date all mitochondria tested have shown VDAC activity. In addition, Table 1 shows that VDAC is not present in rat liver microsomes. The low activity in the microsomal fraction can be accounted for by a 1% mitochondrial contamination. The large standard deviations shown in the table do not indicate irreproducibility in obtaining VDAC activity. Rather, there was a large variation in the number of channels incorporated in the bilayer using the same preparation on the same day. (This type of variability is not unusual when conducting elements are inserted into planar lipid bilayers.) Despite quantitative uncertainty, there is no doubt that VDAC is present in all these organisms.

Figure 1 shows a time course of the current (number of ions crossing the membrane per unit time) which flowed across the membrane after the addition of an extract of rat liver mitochondria to the aqueous phase in which the membrane was bathed. The driving force for ion flow was a 10-mV transmembrane potential. The current increased by distinct jumps of uniform size. The rise time of the jumps was shorter than we can accurately measure, $< 50 \mu\text{s}$ —a characteristic property of channels in general. The rate of channel insertion was fairly constant.

The voltage dependence of VDAC is illustrated in Fig. 2. When the driving force was increased from 10 to 40 mV the current increased fourfold but then decreased to a new lower level. This was due to the switching of VDAC channels from a high to a low conductance state (channels close). When the voltage was returned to 10 mV the channels returned to the high conductance state (channels open). Note that the rate for VDAC opening is much faster than that for closing. The same behaviour

Table 1 Location of VDAC at the organelle level

Source	Major organelle present in fraction	VDAC activity	
		Channels per membrane	Channels per min per mg*
<i>Paramecium</i>	Mitochondria	500 ± 300	
	Cilia	0	
	Pellicle	20 ± 20	
Rat liver	Mitochondria	200 ± 100	
	Microsomes	2 ± 1	0
Rat heart	Mitochondria		330
Beef heart	Mitochondria		370
<i>Neurospora</i>	Mitochondria	700 ± 600	$1,000 \pm 500$
Yeast	Mitochondria	50 ± 50	
Lobster	Axolemma	0	0

Mitochondrial fractions from rat, beef, *Neurospora* (wall-less mutant of *N. crassa*, provided by Dr Gene Scarborough) and yeast were produced by standard procedures. The rat liver microsomes were obtained by centrifuging the post-mitochondrial supernatant at 100,000g for 30 min and taking the top part of the translucent pellet. The *Paramecium* results are those reported previously⁸. Purified lobster axolemma was prepared as described by Barnola *et al.*¹⁶. Channel-forming activity was assayed by two methods. The first was that described previously⁸. [Mitochondria were sonicated with a 25-fold excess of soybean phospholipid, lyophilised and suspended in hexane (1% w/v). The suspension was layered on the aqueous phases to form monolayers which were then used to produce bilayers by the method of Montal and Mueller¹⁷.] Results of this assay are expressed as channels per membrane $\pm$ s.d. In the second method, Triton X-100 was added to the sample to a final concentration of 2% (v/v) (the protein concentration was not greater than 10 mg ml^{-1}). A planar bilayer was made by the method of Montal and Mueller¹⁷ following the technique previously described⁸. (Soybean lipids were used after purification as previously described¹⁸.) An aliquot of this solution (5 or 10 μl) was added to the front compartment. The rate of incorporation of channels into the bilayer, normalised per unit amount of protein added, is a measure of the VDAC activity in the sample (last column). (Note that the Triton extract should be warmed up to room temperature for at least 1 h before assay.) Normalisation is not required in the first assay because the ratio of protein to lipid in the membrane forming solution is kept at a constant 1:25 (w/w). * Rate of VDAC channel insertion from medium per unit time per mg protein.

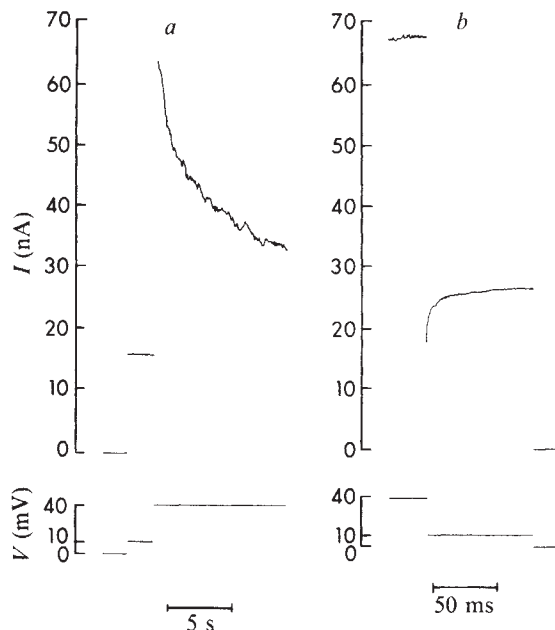


Fig. 2 The voltage-dependent conductance of VDAC. The upper traces show the current and the lower traces the voltage as a function of time. VDAC was inserted as described in Fig. 1. After a certain time the amount of VDAC inserting, for short time periods, is small compared with what has already inserted. Therefore, the conductance can be considered constant for short periods so that conductance changes caused by voltage are not due to channels inserting into the bilayer. On the left-hand graph (a), the voltage was increased from 0 to 10 to 40 mV. The current was constant at 0 and 10 mV. When the voltage was increased to 40 mV, initially, the current increased ohmically indicating no conductance change. Then the current decreased as the VDAC channels closed. On the right-hand graph (b), the membrane was clamped at 40 mV until the current level was fairly constant (channels had turned off). Then the voltage was dropped to 10 mV. The current initially dropped to $\frac{1}{4}$ of its value (ohmic behaviour) and then increased due to VDAC channels opening. Note the difference in time scale between closing and opening rates. The records were first stored using a digital storage oscilloscope (Gould OS 4000) and then transcribed onto a chart recorder.

was seen in VDAC channels from *Paramecium* mitochondria⁸. This was also observed with VDAC channels from *Neurospora crassa* and *Saccharomyces cerevisiae*.

The anion selectivity was verified by measuring the reversal potential of the membrane conductance in the presence of a salt concentration gradient (KCl at 1.0 and 0.1 M). A value of 13 mV positive in the high salt side was measured for material extracted from rat liver—close to that observed in VDAC from *Paramecium* (Fig. 5b of ref. 8).

VDAC's large single channel conductance (0.45 nS in 0.1 M KCl and 4.5 nS in 1.0 M KCl) indicates that the pore size is large. A simplified calculation assuming the specific conductance in the channel to be the same as in solution yields a pore diameter of ~ 15 Å. Large pore size is also suggested by the fact that acetate⁸, sulphate⁸, fumarate and succinate (unpublished observations) easily cross the channel. (Fumarate is five times less permeable than Na^+ .)

VDAC was shown to be located in the outer membrane of rat liver mitochondria by fractionating the mitochondrial membranes into outer and inner (fractions B and P) by the method of Parsons *et al.*⁹. Cross contamination was monitored using cytochrome oxidase activity¹⁰ as a measure of the amount of inner membrane present in the outer membrane fraction and NADH-cytochrome *c* reductase activity¹¹ as a marker for outer membrane. VDAC activity was assayed by following the rate of channel insertion into a lipid bilayer as described above. The results are summarised in Table 2. The marker enzymes confirm good purification of inner and outer membranes with little cross-contamination. VDAC activity is sixfold higher in the

Table 2 Location of VDAC in the mitochondrion

	Inner membrane fraction	Outer membrane fraction	Activity ratio outer/inner
Cytochrome oxidase ($\mu\text{mol cytochrome } c \text{ mg}^{-1} \text{ min}^{-1}$)	1.04	0.063	0.06
NADH-cytochrome <i>c</i> reductase ($\mu\text{mol cytochrome } c \text{ mg}^{-1} \text{ min}^{-1}$)	0.24	6.0	25
VDAC (channels $\text{min}^{-1} \text{ mg}^{-1}$)	$200 \pm 200(7)^*$	$1,200 \pm 1,000(6)$	6

* Mean $\pm$ s.d. with number of assays in parentheses.

outer membrane fraction than in the inner membrane. The large standard deviation is again due to the variability inherent in the assay method. Despite this variability, VDAC is clearly located in the outer mitochondrial membrane. (Its absence from the inner membrane has not been demonstrated.)

Thus the large size of VDAC, its location in the outer mitochondrial membrane and its presence in a wide variety of organisms, indicate that VDAC may be the channel making the outer membrane permeable to small molecules. It is conceivable that the voltage-dependence of VDAC may regulate the flow of metabolites between the cytoplasm and the inner mitochondrial membrane. The small transmembrane potential needed to turn off most of the channels could be generated by a Donnan potential, changes in surface or dipole potentials mediated perhaps by changes in enzyme activity or metabolite concentration.

Interestingly, a similar channel has been isolated from the outer membrane of *Escherichia coli*^{12,13}. Genetic studies¹⁴ and transport measurements of the reconstituted material in vesicles¹⁵ show that there is a channel in the outer membrane of *E. coli* that allows small molecules up to the size of a triose to cross that outer membrane. Benz *et al.*¹² and Schindler *et al.*¹³ have incorporated this protein into planar lipid bilayers and shown that it forms channels that are somewhat smaller than VDAC, weakly selective for cations over anions and voltage dependent, in a similar way to VDAC, but at high transmembrane voltages (150–200 mV compared with 15–30 mV for VDAC). This parallelism may indicate a common primordial channel.

I thank Dr Alan Finkelstein for assistance and advice, and Drs Efraim Racker, Donald Parsons and Carmen Mannella for critical comments. This work was supported by grants from the Scottish Rite Schizophrenia Research Program and the NIH.

M. COLOMBINI

Department Neuroscience and Physiology,
Albert Einstein College of Medicine,
Bronx, New York 10461

Received 8 December 1978; accepted 17 April 1979.

1. Wertheimer, W. C. & Bartley, W. *Biochem. J.* **66**, 79–91 (1957).
2. O'Brien, R. L. & Brierley, G. *J. Biol. Chem.* **240**, 4527–4531 (1965).
3. Pfaff, E., Klingenberg, M., Ritt, E. & Vogell, W. *Eur. J. Biochem.* **5**, 222–232 (1968).
4. Wojtczak, L. & Zaluska, H. *Biochim. biophys. Acta* **193**, 64–79 (1969).
5. Parsons, D. F., Bonner, W. D. Jr & Verboon, J. G. *Can. J. Bot.* **43**, 647–655 (1965).
6. Parsons, D. F., Williams, G. R. & Chance, B. *Ann. N.Y. Acad. Sci.* **137**, 643–666 (1966).
7. Mannella, C. A. & Bonner, W. D. Jr *Biochim. biophys. Acta* **413**, 226–233 (1975).
8. Schein, S. J., Colombini, M. & Finkelstein, A. *J. Membrane Biol.* **30**, 90–120 (1976).
9. Parsons, D. F., Williams, G. R., Thompson, W., Wilson, D. & Chance, B. in *Mitochondrial Structure and Compartmentation* (eds Quagliariello, E., Papa, S., Slater, E. C. & Tager, J. M.) 29–70 (Adriatica Editrice, Bari, 1967).
10. Wharton, D. C. & Tzagoloff, A. *Meth. Enzym.* **10**, 245–250 (1967).
11. Sottocasa, G. L., Kuylenstierna, B., Ernster, L. & Bergstrand, A. *J. Cell Biol.* **32**, 415–438 (1967).
12. Benz, R., Janko, K., Boos, W. & Luger, P. *Biochim. biophys. Acta* **511**, 305–319 (1978).
13. Schindler, H. & Bosenbusch, J. P. *Proc. natn. Acad. Sci. U.S.A.* **75**, 3751–3755 (1978).
14. Bavai, P., Nakaido, H. & von Meyenburg, K. *Molec. gen. Genet.* **158**, 23–33 (1977).
15. Nakae, T. *J. Biol. Chem.* **251**, 2176–2178 (1976); *Biochem. biophys. Res. Commun.* **71**, 877–889 (1976).
16. Barnola, F. V., Villegas, R. & Camejo, G. *Biochim. biophys. Acta* **298**, 84–94 (1973).
17. Montal, M. & Mueller, P. *Proc. natn. Acad. Sci. U.S.A.* **69**, 3561–3566 (1972).
18. Kagawa, Y. & Racker, E. *J. Biol. Chem.* **246**, 5477–5487 (1971).

A stable chemiluminescent-labelled antibody for immunological assays

RADIOISOTOPES, in particular ^{125}I , have been used for many years as labels for monitoring the distribution of reagents in immunological assay systems. Satisfactory labelled derivatives are not always easily produced as iodination frequently produces severe molecular disruption. Moreover, iodinated proteins have a limited shelf life and sometimes require extended counting times for accurate quantitation. Recently, non-radioactive labels such as bacteriophages^{1,2}, enzymes^{3,4}, stable free radicals⁵ and fluorescent groups⁶ have been used in attempts to overcome some of these problems. We describe here the successful labelling of antibodies to rabbit IgG with a chemiluminescent molecule and the development of an immunological assay using chemiluminescence as a means of monitoring reagent distribution.

The luminescence of luminol (5-amino-2,3-dihydrophthalazin-1,4-dione) occurs during oxidation in alkaline conditions⁷ in the presence of certain inorganic ions or molecules⁸ (for example MnO_4^- , I_2 , $\text{Fe}(\text{CN})_6^{3-}$, Cu^{2+} and OCl^-). In such conditions, luminol takes part in a reaction which generates an excited product molecule. This subsequently becomes de-excited by several mechanisms⁹, one of which is the emission of photons. The quantum yield for this process is approximately 0.015 in aqueous conditions¹⁰. A photon counter built in our laboratory¹¹ is capable of detecting 10^{-18} mol of luminol.

An IgG fraction of a sheep anti-rabbit IgG antiserum prepared by sodium sulphate precipitation was reacted with a diazonium salt of luminol. Diazotisation involved a marked reduction in quantum yield (Table 1), possibly as a result of polymerisation of diazoluminol. A further small reduction took place following overnight incubation at pH 8.6 as required for the coupling reaction, although there was no further loss of quantum yield as a result of the coupling procedure itself. The incorporation of diazoluminol into conjugate with IgG was dependent on the concentration of reactants (Table 2). Using a 100-fold molar excess of luminol it was possible to incorporate up to 80% of the ligand, although it was found that at uptake

Table 1 Recovery of luminescence during coupling of luminol to IgG

Stage of coupling	Total recovered counts	% Initial luminescence
Initial	5.9×10^{11}	100
Diazoluminol*	7.9×10^9	1.34
Diazoluminol†	3.8×10^9	0.67
Lumino-IgG conjugate	4.0×10^9	0.70

Luminol-IgG conjugate was prepared by diazotising 200 mg of luminol in 20 ml ice-cold 2.4 M HCl with 200 mg NaNO_2 . Ice-cold borate buffer (50 ml 0.5 M, pH 8.5) was added and the pH adjusted to 8.5. Fourteen ml of this solution was added to 8.5 ml of a solution containing 1 g of an IgG fraction, prepared by sodium sulphate precipitation of a sheep anti-rabbit IgG antiserum (code SB 82/2) in borate buffer (0.5 M, pH 8.5). The molar ratio of luminol: IgG was calculated to be approximately 30:1. The mixture was allowed to stand for 18 h at 4 °C to enable coupling to occur. Unconjugated luminol was removed by extensive dialysis against 0.2 M borate buffer pH 8.0 and stored as 0.5 ml aliquots at -20 °C. Samples (10 μl) were removed at stages during the preparation of the conjugate and diluted (1/100–1/1,000) with 0.1 M NaOH. Luminescence was measured as follows. Aliquots (10 μl) of the diluted samples were added to 1 ml of 0.1 M NaOH in polystyrene tubes. 5.4 mM H_2O_2 (100 μl) was added and after mixing the reaction tube was placed in the reaction chamber of the digital photon counter. The reaction was initiated by addition of 1 ml 1.0 mM NaOCl in 0.1 M NaOH solution. Counts were accumulated for a period of 20 s. After correction for volume changes values were calculated to give total recovered counts.

* Immediately on preparation.

† After overnight incubation at 4 °C in 0.2 M borate buffer (pH 8.5).

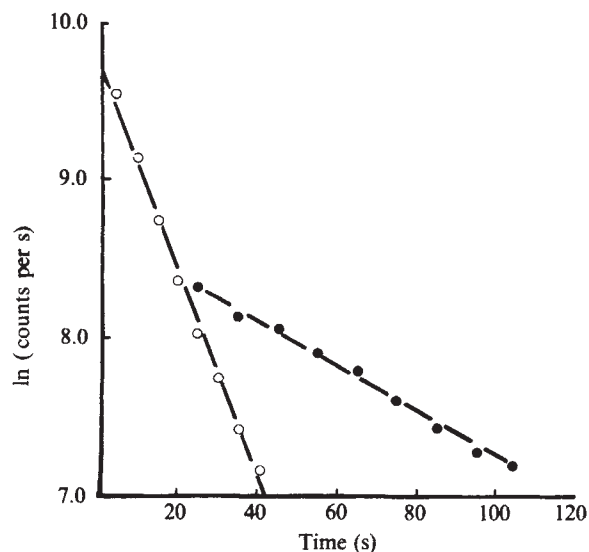


Fig. 1 Time course of light emission from luminol (●) and luminol-IgG conjugate (○). The luminol-IgG conjugate was prepared as described in Table 1, and purified before use by affinity chromatography using an immunoadsorbent (ImAd) prepared from rabbit IgG coupled to a diazonium salt of powdered cellulose¹². After reaction overnight at 4 °C the ImAd-antibody complex was washed free of unbound label using 0.05 M veronal buffer, pH 8.2 containing (w/v) 0.1% bovine serum albumin, 0.05% sodium azide 0.002% non-immune sheep IgG (NISH buffer). After washing with water to remove the buffer, luminol-labelled antibody was eluted by reduction of the pH to 2.0. The supernatant was buffered to pH 8.2 using an equal volume of NISH buffer at double strength with respect to all constituents. The luminescence was quantitated as described in Table 1. In this case the rate of light emission was plotted on a chart recorder. Values taken from this graph were plotted as the natural logarithm against time (in seconds) from the start of the reaction.

ratios in excess of 40 mol luminol per mol IgG the product precipitated on storage. The labelled product was freed of unreacted luminol by extensive dialysis and purified subsequently by affinity chromatography using a rabbit IgG immunoabsorbent. The purified product prepared by reacting IgG of the sheep anti-rabbit IgG antiserum with a 30-fold molar excess of diazoluminol contained 22 mol luminol per mol IgG.

Figure 1 compares the time courses of light emission from luminol and the luminol-antibody conjugate. The two materials were tested in identical conditions as the rate constant for light emission is dependent on the concentration of H_2O_2 and catalyst⁸. Luminol conjugated to IgG had a significantly higher rate constant than luminol itself. When the luminescent material was the limiting reactant, the decay phase of the light-emitting reaction was an exponential curve with a rate constant of 0.073 s^{-1} for conjugated luminol and a value of 0.014 s^{-1} for luminol itself. The reason for the increase in reaction rate of conjugated luminol is not yet clear. It is possible that some structural alteration in the ligand has taken place.

In these experiments we have quantified the luminescent material by measuring accumulation of counts during a 20-s period. It is apparent from Fig. 1, however, that in the conditions used for these experiments, the rate of emission of light from the reaction is first order with respect to the luminol concentration. It is therefore possible to quantify the luminescent material by measuring the maximum reaction rate over a short period, for example 2 s.

The immunological reactivity of the luminol-labelled sheep anti-rabbit IgG antibody was assessed by a radioimmunoassay using solid phase rabbit IgG and ^{125}I -labelled sheep anti-rabbit IgG antibody. Through a series of dilutions (1/1 to 1/1,000 with respect to the antiserum) no difference could be detected between the luminol-labelled antibody and unlabelled sheep anti-rabbit IgG antibody in ability to bind antigen.

Using the luminol-labelled sheep anti-rabbit IgG antibodies it was possible to develop a two-site assay¹³ for rabbit IgG. The procedure involved extraction of rabbit IgG on to polystyrene tubes coated with unlabelled sheep anti-rabbit IgG antibody. Uptake of the antigen was then determined by reacting with the labelled antibodies. Figure 2 shows standard curves obtained using luminol-labelled antibodies immediately after preparation (a) and after 9 months storage at -20°C (b) with a curve obtained with freshly prepared ^{125}I -labelled sheep anti-rabbit IgG antibody. The luminol-labelled antibody produced an assay of similar sensitivity to that obtained with the iodinated antibody but also showed no apparent loss of activity after storage for 9 months.

These results show that antibody molecules may be labelled with luminol without altering their immunological reactivity, and, despite losses in the quantum yield from the luminescent molecule, these antibodies may be used to develop a satisfactory immunological assay. Moreover, the time course of light emission from the labelled derivative is rapid enough for accurate measurement to be made within 2 s of the initiation of the luminescent reaction. Finally, the stability of the labelled antibody is such that assays may be carried out over long periods (at least 9 months) with no apparent loss of performance.

The development of chemiluminescent labels offers considerable potential for improvements in immunoassay technology and in other areas of biochemistry and histology where there is a need for biologically active high specific activity labels. Because they are stable, such reagents can be produced in bulk. This has important implications in terms of cost and quality control of assays. The ability to carry out extremely rapid

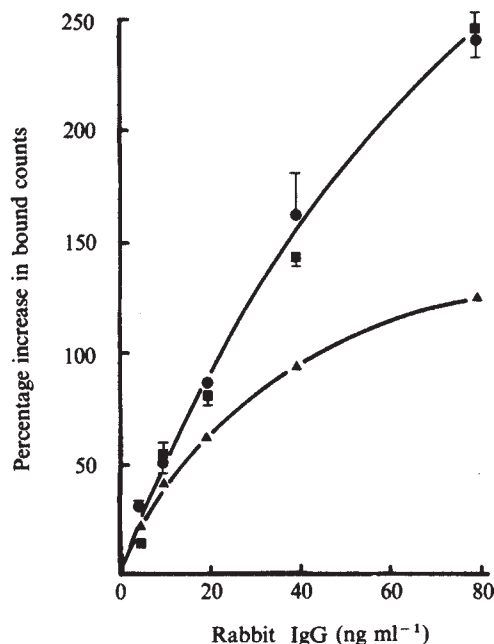


Fig. 2 Standard curves obtained in a two-site assay of rabbit IgG using luminol-labelled and ^{125}I -labelled (▲) sheep anti-rabbit IgG antibody. Assays were carried out with freshly prepared luminol-labelled antibody (●) and the same reagent after nine months storage at -20°C (■). Antibodies purified from antiserum SB 82/2 by means of affinity chromatography with rabbit IgG ImAd, were incubated in polystyrene tubes in 0.1 M carbonate buffer pH 9.6 at a dilution equivalent to 1:85 of the original serum, to prepare a solid phase derivative¹⁴. After coating, the tubes were washed with 0.1% bovine serum albumin in 0.15 M sodium chloride solution containing 0.1% thiomersal, and stored empty at -20°C . In the assay procedure, standard solutions of rabbit IgG in NISH buffer (half strength) were added to the tubes and incubated overnight at 4°C . Then, either luminol or ^{125}I -labelled antibodies purified as described in Fig. 1 were added. After a further overnight incubation at 4°C , the tubes were washed again in NISH buffer (half strength) and either counted in an automatic gamma spectrometer or assayed for luminescence activity as described in Table 1.

Table 2 Incorporation of luminol into luminol-IgG conjugate

Initial luminol/IgG ratio	A_{450}^a	A_{450}^b	TCA precipitated ^c (%)	Molar ratio of luminol: IgG ^d
0	0.000	0.000		
5	0.172	0.098	43.0	2.15
10	0.199	0.067	66.3	6.63
20	0.228	0.074	67.5	13.5
50	0.419	0.077	81.6	40.8
100	0.758	0.127	83.25	83.25

Luminol and diazoluminol are both soluble in TCA solutions. Both compounds absorb strongly at 450 nm.

^a Absorbance (450 nm) of conjugate after coupling, before dialysis.

^b Absorbance (450 nm) after precipitation of protein by addition of an equal volume of 15% trichloroacetic acid (TCA). The supernatant was buffered to pH 8.5 and a correction made for volume change.

^c $(1 - b/a) \times 100\%$.

^d $c/100 \times \text{initial ratio}$.

quantitation of label is likely to prove an important factor in the organisation of assays involving large sample numbers, for example screening programmes. Finally, the quantitation of chemiluminescence offers a major advantage, particularly with respect to sensitivity, over other non-isotopic labels (for example fluorescence) in that detection is made against a background which is theoretically zero.

We acknowledge an equipment grant from the SRC. J.S.A.S. was in receipt of an MRC research scholarship.

J. S. A. SIMPSON
A. K. CAMPBELL
M. E. T. RYALL
J. S. WOODHEAD

Department of Medical Biochemistry,
Welsh National School of Medicine,
Cardiff, UK

Received 5 February; accepted 27 April 1979.

- Haimovich, J., Hurwitz, E., Norvik, N. & Sela, M. *Biochim. biophys. Acta* **207**, 125-129 (1970).
- Andrieu, J.-M., Mamas, S. & Dray, F. in *Steroid Immunoassay* (eds Cameron, E. H. D., Hillier, S. G. & Griffiths, K.) 189-198 (Alpha Omega, Cardiff, 1975).
- Engvall, E. & Perlmann, P. *Immunochemistry* **8**, 871-874 (1971).
- Van Weemen, B. K. & Schuur, A. H. W. M. *FEBS Lett.* **15**, 232-236 (1972).
- Leute, R., Ullman, E. F. & Goldstein, A. *J. Am. med. Ass.* **221**, 1231-1234 (1972).
- Aalberse, R. C. *Clin. chim. Acta* **48**, 109-111 (1973).
- McCapra, F. Q. *Rev. Chem. Soc.* **20**, 485-510 (1966).
- White, E. H. & Brendett, R. B. in *Chemiluminescence and Bioluminescence* (eds Cormier, M. J., Hercules, O. M. & Lee, J.) 231-244 (Plenum, New York, 1973).
- Gundermann, K.-D. *Curr. Topics Chem.* **46**, 61-139 (1974).
- Nakamura, T. *J. Biochem.* **72**, 173-177 (1972).
- Campbell, A. K. & Dormer, R. L. *Biochem. J.* **152**, 255-265 (1975).
- Gurevich, A. E., Kuzovleva, O. B. & Tumanova, A. E. *Biokhimiya* **26**, 803-809 (1962).
- Woodhead, J. S., Addison, G. M. & Hales, C. N. *Br. med. Bull.* **30**, 44-49 (1974).
- Catt, K. & Tregear, G. W. *Science* **158**, 1570-1572 (1967).

Introns in the chicken ovalbumin gene prevent ovalbumin synthesis in *E. coli* K12

IT has recently been discovered that some eukaryotic genes are split, the mRNA coding sequences (exons) being interrupted by intervening sequences (introns) of unknown function¹⁻³. In all cases examined, these split genes are transcribed into an RNA precursor which is then matured by a splicing mechanism⁴⁻⁶. Because this type of gene structure has not previously been found in bacteria, it is generally believed that split genes cannot be expressed in *Escherichia*, but there are no facts to support this assumption. We demonstrate here that the presence of introns in the chicken ovalbumin gene prevents ovalbumin synthesis in *E. coli* K12.

The chicken ovalbumin gene is split into eight exons (designated 1 to 8) separated by seven introns (named A to G)^{2,7-17}. It

is transcribed into an RNA precursor (at least) 7,600 nucleotides long, which is then spliced into ovalbumin mRNA (ov mRNA)^{5,6}. We have isolated, by molecular cloning in *E. coli* K12, chicken DNA segments which carry part of the gene or the entire gene⁷⁻¹⁴. We have also constructed plasmids which carry synthetic double-stranded DNA sequences complementary to ov mRNA (ov ds cDNA) and are devoid of introns¹⁷. In one of these plasmids, pOMP2, the ovalbumin sequence was fused in phase at the beginning of the *E. coli* gene coding for β -galactosidase, the fifth amino acid of ovalbumin (from the NH₂-terminal part of the protein) being connected to the seventh amino acid of β -galactosidase¹⁸. We have shown (as have others¹⁹) that *E. coli* cells harbouring the recombinant plasmid synthesise large amounts of an ovalbumin-like product, which resembles ovalbumin in size and reactivity with ovalbumin-specific antibodies. We have manipulated the fused *lac*-ovalbumin sequence and introduced into it some of the introns present in the natural gene so as to assess their influence on the synthesis of the ovalbumin-like product.

There are two sites for the restriction endonuclease *Sst*I in ov ds cDNA, 378 base pairs apart, located at nucleotides 116 and 494 in the ov mRNA sequence^{8,11}. In the natural gene, these sites are located in exons 1 and 4, respectively, and separated by 1,612 base pairs which include introns B, C and D (refs 8, 11). The substitution of this 1,612-base pair fragment in place of the 378-base pair one in the *lac*-ovalbumin sequence will thus result in the introduction of three introns of the natural gene. This was achieved as depicted in Fig. 1: because pOMP2 has an additional *Sst*I site (in a sequence originating from phage λ (ref.

18)) we first extracted the *lac* regulatory region with the *lac*-ovalbumin fusion by digestion with *Hha*I. The resulting 2.67-kilobase fragment was introduced into partially digested pBR322, cut only once by *Hha*I. The resulting plasmids, like pOMP2, produced large amounts of ovalbumin-like protein in *E. coli*. One of these, pOMP21, was selected for further studies. The 378-base pair *Sst*I fragment of pOMP21 was substituted by the 1,612-base pair *Sst*I fragment extracted from the cloned *Eco*RI fragment 'b' of the ovalbumin gene⁸. We thus obtained pOMP23 in which introns B, C and D were correctly orientated for transcription from the *lac* promoter (see Fig. 1 legend for mapping of this plasmid). *E. coli* harbouring pOMP23 was then tested for the production of ovalbumin-like protein by a radioimmunoassay. As shown in Fig. 2, pOMP21 (as pOMP2) makes large amounts of an ovalbumin-like product, whereas no ovalbumin-like product is detectable in the pOMP23 assay, even with concentrated bacterial extracts—at least 3,000-fold less than pOMP21, that is, less than three molecules per bacterium in the experiment shown.

We next investigated whether this negative result could be due to inadvertent mutational changes during the construction of pOMP23. The *Sst*I 1,612-base pair fragment was excised from pOMP23 and replaced by the 378-base pair *Sst*I fragment extracted from ov ds cDNA. Where the insertion occurred in the proper direction, pOMP26 again produced large amounts of ovalbumin-like protein (not shown). This does not exclude the possibility that mutations occurred in the 1,612-base pair fragment. The latter, however, was isolated from the same cloned *Eco*RI 'b' fragment of the ovalbumin gene which was sequenced

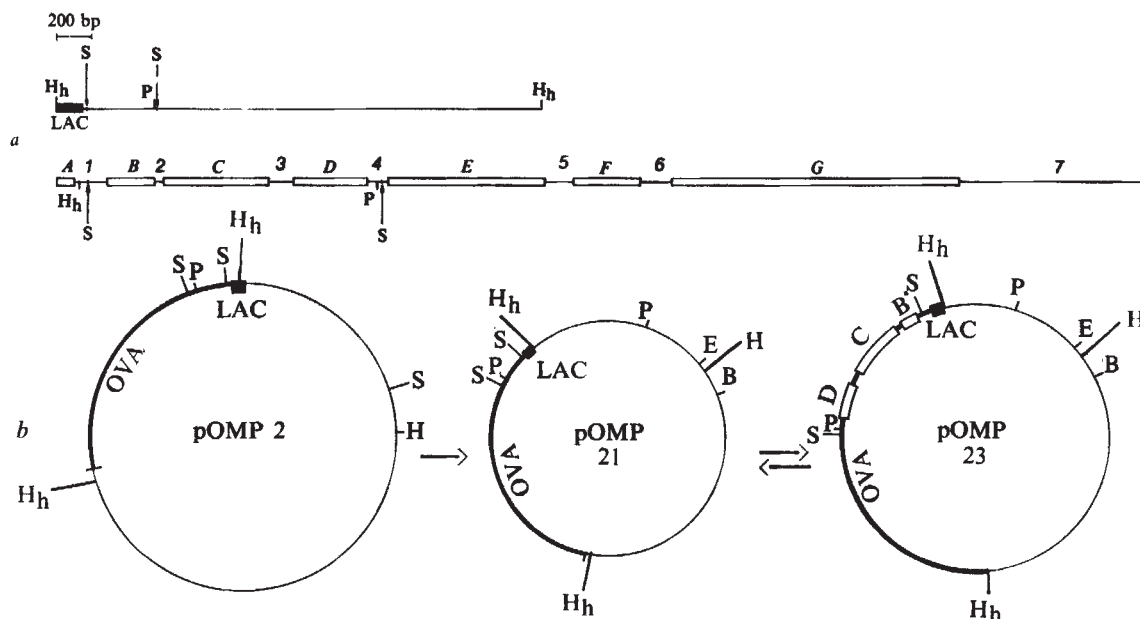


Fig. 1 Introduction of introns B, C and D into the *lac*-ovalbumin sequence. *a*, Restriction maps of the fused *lac*-ov ds cDNA sequence and of the ovalbumin *Eco*RI 'b' fragment (see also refs 11, 14, 15). The *Hha*I site in the *lac* regulatory region lies immediately after the TGA stop signal at the end of the *lac* I gene². The *Hha*I site in exon 1 is the one used to construct the *lac*-ovalbumin in fused sequence in pOMP2 (refs 8, 18). *b*, Map of the plasmids pOMP2, pOMP21 and pOMP23, pOMP2 and its related plasmids have 32 *Hha*I sites²¹. Only those used to excise the *lac*-ovalbumin sequence are shown. The construction of pOMP23 involved the following steps: 20 μ g of pOMP2 DNA were digested to completion by *Hha*I (Biolabs; 25 units; 30 min at 37 °C) and the reaction was stopped by heating to 65 °C for 10 min in the presence of 10 mM EDTA. The reaction mixture was loaded on a 5–20% sucrose gradient (w/v in 25 mM Na acetate, pH 6.0, 10 mM EDTA, 0.5% SDS) and centrifuged at 31,000 r.p.m. for 16 h at 20 °C in a SW41 rotor of a Beckman ultracentrifuge. The 2.67-kilobase fragment carrying the fused *lac*-ovalbumin sequence could thus be purified from the other, much smaller fragments of pOMP2. pBR322 (ref. 21) DNA (20 μ g) was partially digested by *Hha*I. Enough enzyme was added to convert about half of the supercoils into linear molecules, which were purified by centrifugation through a 5–20% sucrose gradient (w/v in 10 mM Tris-HCl, pH 8.0, 100 mM NaCl, 1 mM EDTA; 4 h at 40,000 r.p.m. and 20 °C in a SW41 Beckman rotor). 200 ng of the 2.67-kilobase *lac*-ovalbumin *Hha*I fragment and 150 ng of linearised pBR322 were ligated with T4 DNA ligase at 4 °C in a final volume of 5 μ l in 40 mM Tris, pH 7.5, 10 mM MgCl₂, 10 mM dithiothreitol, 0.1 mM ATP, 0.05 mg ml⁻¹ bovine serum albumin. After transformation of CaCl₂-treated lysogenic cells of strain 1398 (ref. 22), transformants were selected on ampicillin-containing plates (20 μ g ml⁻¹) and screened by *in situ* hybridisation²² with a ³²P-labelled probe specific for the ovalbumin sequence (the *Hha*ov probe^{7,8}). Positive colonies were then tested for ovalbumin production in the same radioimmunoassay as before¹⁸. All the clones positive in *in situ* hybridisation were also positive in this assay. Plasmid pOMP21 was arbitrarily chosen for further studies, and transferred into the non-lysogenic strain C600rk⁻mk⁺. pOMP21 (5 μ g) was then digested by *Sst*I (BRL, 2.5 h at 37 °C) and centrifuged in a 5–20% sucrose gradient (v/v in 10 mM Tris-HCl, pH 7.4; 20 mM NaCl; 1 mM EDTA; 45,000 r.p.m. for 140 min at 20 °C in a SW 50.1 Beckman rotor). The 6.65-kilobase fragment was treated with calf intestine alkaline phosphatase²²; of this, 133 ng was ligated by T4 DNA ligase in a final volume of 5 μ l, with 100 ng of the 1.61-kilobase fragment released after digestion by *Sst*I of pBR322 carrying the *Eco*RI 'b' fragment of the ovalbumin gene. After transformation of C600rk⁻mk⁺, 12 ampicillin-resistant clones were analysed. Purified clear lysates were treated by *Hha*I and examined for the presence of the 3.9-kilobase fragment containing introns B, C and D. In such clones, the orientation of the integrated 1.61-kilobase fragment was determined by digestion with *Pst*I and *Hha*I. The unique *Pst*I site in the ovalbumin sequence lies at nucleotide 477 in exon 4 close to one end of the integrated fragment. In such analyses, plasmid pOMP23 yielded fragments of 2.13 and 1.78 kilobases (instead of 0.19 and 3.72 kilobases in the reverse orientation), demonstrating proper orientation of introns B, C and D. More precise mapping of plasmids pOMP21 and pOMP23 was achieved by the use of *Pst*I, *Bam*HI and *Sac*I. It indicated that the *lac*-ovalbumin fused sequence was most probably integrated into the *Hha*I site located at nucleotide 2,351 of the pBR322 sequence²⁴. Replacement of the 1,612-base pair fragment of pOMP23 by the 378-base pair ov ds cDNA segment was carried out by the same procedures as above, except that *Sac*I was used instead of its isoschizomer *Sst*I (ref. 25). The letters S, P, E, B, H and Hh refer to *Sst*I, *Pst*I, *Eco*RI, *Bam*HI, *Hind*III and *Hha*I sites, respectively.

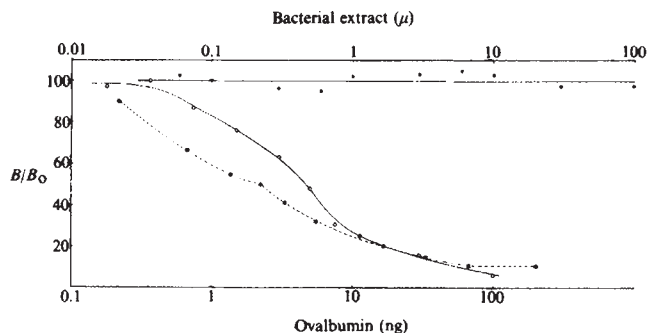


Fig. 2 Synthesis of ovalbumin-like protein by pOMP21 and pOMP23. C600rk⁺mk⁺ (pOMP21) and C600rk⁺mk⁺ (pOMP23) were grown overnight in L broth supplemented with 20 $\mu\text{g ml}^{-1}$ ampicillin, diluted into fresh medium containing 1 mM isopropylthiogalactoside, grown to about 4×10^8 cells per ml, concentrated about 100-fold, and sonicated twice for 30 s as before¹⁸. The procedures for iodination of ovalbumin, purification of antibodies and radioimmunoassays have been described previously¹⁸. In the experiment shown, the dilution of antiserum chosen was able to precipitate 30% of the iodinated ovalbumin (B_0). Results are expressed as B/B_0 , the ratio between the c.p.m. precipitated in the presence or the absence of competing material (blanks being subtracted each time). B was measured in duplicate and B_0 was the average of four determinations. ○, Results obtained with pure ovalbumin; *, those obtained with pOMP21, representing the synthesis of about 6,500 molecules per bacterium. ●, Results obtained with pOMP23. Similar results are obtained in various hosts. C600rk⁺mk⁺ carries the *supE* suppressor for amber mutations.

by K. O'Hare *et al.* (personal communication), and where the sequence of the various exons matched those of *ov* mRNA. It thus seems unlikely that inadvertent mutations cause the defect in expression of the ovalbumin-like product. Instead, we conclude that the introduction of introns *B*, *C* and *D* in the fused *lac*-ovalbumin sequence blocks the expression of ovalbumin-like product in *E. coli*.

The sequence data of K. O'Hare *et al.* indicate that introns *B* and *C* are 252 and 582 base pairs long (84 and 194 triplets, respectively) so that their presence will not alter the reading frame. In contrast, intron *D* consists of 401 base pairs (133 triplets plus 2 base pairs) and will cause a frameshift. (The sequence data of Robertson *et al.*²⁸, obtained with another cloned DNA, show some differences; in their case, intron *C* would also cause a frameshift.) Further, intron *B* contains one nonsense codon (TGA), intron *C* shows 12 nonsense codons (four TAA, three TGA and five TAG) and intron *D* contains three TGA and three TAA stop signals. All these nonsense codons are in the same reading frame, and although the bacterial host used had some suppressing capabilities (see legend to Fig. 2) it is unlikely that efficient suppression could occur. Because, in addition, intron *D* causes a frameshift, any synthesis of ovalbumin-like product in *E. coli* would probably result from the splicing of intron sequences from the primary RNA transcript. The mechanisms involved in RNA splicing in the higher eukaryotes are still poorly understood. It is possible that additional, unknown eukaryotic genetic elements could promote splicing in *E. coli*. Also, we cannot rule out the possibility that the simultaneous presence of the seven introns of the ovalbumin gene is a prerequisite for splicing in *E. coli*. It seems very likely, however, that splicing does not readily occur in *E. coli*, and that the isolation of split genes or the production of the corresponding proteins cannot be based at this stage on their expression in *E. coli*.

Biohazards associated with the experiments described here were examined by the French National Control Committee. We thank Dr P. Chambon, K. O'Hare, C. Benoist and R. Breathnach for communicating unpublished results (especially sequence data) and for interesting discussion, Dr F. Rougeon for enzymes, Mrs O. Le Bail for technical help and Professor Fauve for the use of his γ counter. This work was supported by grants from CNRS (ERA No. 201 and ATP No. 3558), INSERM (CRAT No. 76.4.310 and 76.4.465), FRMF and DGRST.

O. MERCEREAU-PUJALON
P. KOURILSKY

Groupe de Biologie Moléculaire du Gène,
ERA CNRS 201 Unité de Génie Génétique,
Institut Pasteur,
28, Rue du Docteur Roux,
75724 Paris Cedex 15, France

Received 9 April; accepted 14 May 1979.

1. Jeffreys, A. J. & Flavell, R. A. *Cell* **12**, 1097-1108 (1977).
2. Breathnach, R., Mandel, J. L. & Chambon, P. *Nature* **270**, 314-319 (1977).
3. Leder, P. *et al.* *Cold Spring Harb. Symp. quant. Biol.* **42**, 915-920 (1977).
4. Tilghman, S. *et al.* *Proc. natn. Acad. Sci. U.S.A.* **75**, 1309-1313 (1978).
5. Roop, D. R. *et al.* *Cell* **15**, 671-685 (1978).
6. Chambon, P. *et al.* *Proc. 11th Winter Miami Symp.* (in the press).
7. Garapin, A. *et al.* *Nature* **273**, 349-354 (1978).
8. Garapin, A. *et al.* *Cell* **14**, 629-639 (1978).
9. Mandel, J. L. *et al.* *Cell* **14**, 641-653 (1978).
10. Le Penne, J. P. *et al.* *Nucleic Acids Res.* **5**, 4547-4562 (1978).
11. Breathnach, R. *et al.* *Proc. Natn. Acad. Sci. U.S.A.* **75**, 4853-4857 (1978).
12. Kourilsky, P. & Chambon, P. *Trends biochem. Sci.* **3**, 244-247 (1978).
13. Perrin, F. *et al.* in *Genetic Engineering* (eds Boyer H. W. & Nicosia, S.) 89-97 (Elsevier, Amsterdam 1978).
14. Gannon, F. *et al.* *Nature* **278**, 428-434 (1979).
15. McKeaynolds, L. *et al.* *Nature* **273**, 723-728 (1978).
16. Dugaiczky, A. *et al.* *Nature* **274**, 328-333 (1978).
17. Humphries, P. *et al.* *Nucleic Acids Res.* **4**, 2389-2406 (1977).
18. Mercereau-Pujalon, O. *et al.* *Nature* **275**, 505-510 (1978).
19. Fraser, T. H. & Bruce, B. *Proc. natn. Acad. Sci. U.S.A.* **75**, 5936-5940 (1978).
20. Dickson, R. *et al.* *Science* **187**, 27-35 (1975).
21. Bolivar, F. *et al.* *Gene* **2**, 95-113 (1977).
22. Cami, B. & Kourilsky, P. *Nucleic Acids Res.* **5**, 2381-2390 (1978).
23. Sippel, A. E. *et al.* *Nucleic Acids Res.* **5**, 3275-3294 (1978).
24. Sutcliffe, J. G. *Nucleic Acids Res.* **5**, 2721-2728 (1978).
25. Roberts, R. J. *Gene* **4**, 183-193 (1978).
26. Robertson, M. A. *et al.* *Nature* **278**, 370-371 (1979).

About 30% of minute virus of mice RNA is spliced out following polyadenylation

MANY viral and eukaryotic messenger RNAs have been shown to contain blocks of sequence which, although adjacent to one another in the RNA, are copied from non-contiguous regions of the genome that are joined together post-transcriptionally. This phenomenon of RNA splicing has been shown to occur during the processing of several DNA and RNA tumour viruses¹⁻⁷ and cellular⁸⁻¹⁰ mRNAs. The mechanism by which splicing is carried out has not been elucidated, and any direct approach is complicated by the elaborate splicing patterns found in the systems which have been examined so far. We have extended these studies of RNA splicing to a potentially simpler viral system, the minute virus of mice (MVM), a non-defective member of the parvovirus group. The members of this group are considered to be among the simplest of the animal viruses, and comprise an icosahedral virion made from three structural protein species, containing one linear single-stranded DNA chromosome of $\sim 1.5 \times 10^6$ molecular weight (for review see ref. 11). The MVM DNA molecule contains a stable hairpin duplex at its 5'-end and a 3'-terminal structure suitable for priming complementary strand synthesis *in vitro*^{12,13}. Little is known about the transcription process in parvoviruses. The only cytoplasmic mRNA species found for adeno-associated virus (AAV), a member of the defective subgroup, is a transcript of 70% of the genome which has been shown by *in vitro* translation to code for the three capsid proteins¹⁴. The one study of transcription carried out in the non-defective subgroup has identified an analogous RNA species in Kilham rat virus (KRV)-infected cell cytoplasm which is complementary to 70% of the viral strand and sediments as a single component at about 18S (ref. 15). We report here our attempts to determine the sequence arrangement of MVM RNA by examining, in the electron microscope, hybrids formed between the abundant species of nuclear and cytoplasmic poly(A)⁺ RNA and single-stranded genomic DNA. We have found that about 30% of MVM RNA is spliced out following polyadenylation.

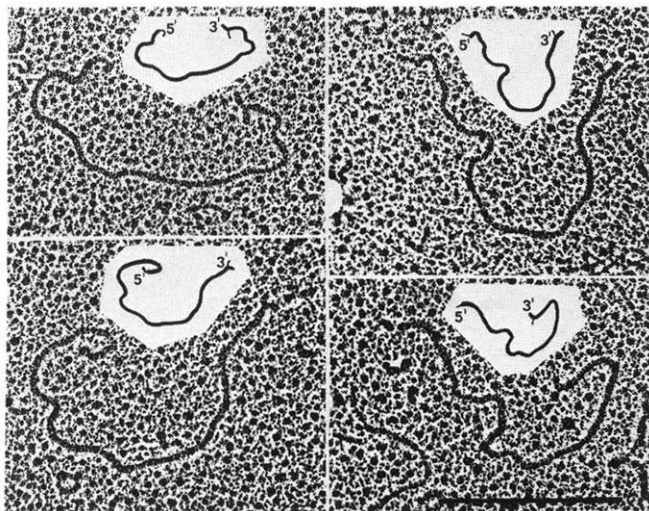


Fig. 1 Electron micrographs and their tracings of hybrids formed between nuclear poly(A)⁺ RNA and MVM DNA. A-9 cells, a derivative of the mouse L-cell line (3×10^7 cells) were infected with 1–2 plaque-forming units per cell of plaque-purified strain T of MVM. At 20 h p.i., nuclear and cytoplasmic fractions were prepared¹⁸, and RNA extracted¹⁹ and fractionated on oligo(dT) cellulose columns to poly(A)⁺ and poly(A)[−] molecules. Genomic single-stranded DNA was extracted from purified virions¹². The RNA was precipitated in ethanol with the genomic single-stranded DNA (0.5 µg). The pellet was collected by centrifugation and resuspended in 66% (v/v) of repurified formamide, 3 M urea, 20 mM EDTA and 0.2 M Tricine-NaOH, pH 8.1. The hybridisation mixture was incubated at 37 °C for 6 h. Samples were prepared for visualisation in the electron microscope as detailed by Bratosin *et al.*²². Scale bar, 0.5 µm.

Hybrids were obtained by annealing poly(A)⁺ nuclear RNA and MVM genomic DNA. Figure 1 shows the appearance of representative RNA–DNA structures with their schematic tracings. Molecules of unit length MVM DNA were selected and examined for single- and double-stranded regions. Of more than 100 molecules scored, at least 95% exhibited double-stranded appearance over the entire length of the DNA. It should be emphasised, however, that the present analysis would not detect single-stranded regions of less than 50 nucleotides. Each hybrid examined shows a short Y-shaped structure at one end, consisting of one double-stranded and one single-stranded arm. The lengths of the two arms support the assumption that the double-stranded arm is the 130-base pair hairpin known to be at the 5'-end of the viral DNA¹² and that the single-stranded arm represents at least in part the poly(A) on the 3'-end of the RNA. The opposite end of these hybrid structures exhibit, in 50% of molecules examined, a short single-stranded region of up to 150 nucleotides, suggesting that the initiation of transcription might occur in a number of alternative sites. In this context it is noteworthy that SV40 nuclear and cytoplasmic RNAs have various 5'-terminal sequences^{20,21}. As the complete (or almost complete) double-stranded hybrids formed with nuclear RNA are by far the most abundant structures observed in the electron microscope, it seems that poly(A)⁺ nuclear RNA is a complete (or almost complete) transcript of the genomic DNA, and suggests that it is the precursor for cytoplasmic poly(A)⁺ mRNA. Moreover, because the template for transcription is at least in part a linear DNA molecule (unpublished results), we conclude that transcription is always initiated within a region located at the extreme 3'-end of the viral chromosome. This correlates well with the proposed structure of the nuclear precursor for the AAV messenger RNA¹⁴.

RNA was extracted from the cytoplasmic fraction of MVM-infected cells at 20 h post-infection (p.i.). Poly(A)⁺ RNA was selected by chromatography on oligo(dT) cellulose columns. About 25% of ³H-uridine-labelled poly(A)⁺ cytoplasmic RNA (labelled 16–20 h p.i.) hybridised with genomic DNA, indicating

that the viral RNA comprises a substantial fraction of the poly(A)⁺ RNA present in the infected cells.

Hybrids were obtained by annealing the poly(A)⁺ cytoplasmic RNA and MVM genomic DNA. Figure 2 shows the appearance of representative RNA–DNA structures with their schematic tracings. As in the analysis of nuclear RNA, hybrids containing unit length MVM DNA were selected and examined for single- and double-stranded regions. Of about 100 molecules scored, more than 85% exhibited one short and one long arm of duplex with an intervening single-stranded DNA loop. As with the hybrids containing nuclear RNA, one end, that of the long duplex arm, has a short Y-structure comprising one double strand and one single strand, which we again postulate is the 3'-end of the RNA. As with the nuclear RNA–DNA hybrids, the opposite end, that of the shorter duplex arm, exhibited a short single-stranded region in many of the molecules examined. The observation made with the nuclear and cytoplasmic poly(A)⁺ viral RNAs suggests that the MVM poly(A)⁺ cytoplasmic RNA is a spliced molecule composed of two parts: a short 'leader' and a longer 'body'. The single-stranded DNA loop (see Fig. 2) contains the intervening template for RNA sequences which are present in the nuclear poly(A)⁺ RNA and have been spliced out in the cytoplasmic poly(A)⁺ RNA. The analysis of nuclear and cytoplasmic RNAs also indicates that the sequence of modifications during mRNA maturation is polyadenylation followed by splicing. We have measured the three parts on the RNA–DNA structures shown in Fig. 2 and the results of these measurements are presented in Fig. 3. Sharp histograms were obtained for each of the three parts: the leader spans $8.8 \pm 1.2\%$ of the structure, the single-stranded DNA loop (or intron) spans $29.3 \pm 2.4\%$ of the structure and the body spans $61.8 \pm 2.3\%$ of the structure. The map coordinates of each part of the structures are represented in Fig. 3b.

The present study does not indicate whether the poly(A)⁺ cytoplasmic RNA which we have analysed is the only viral RNA present in the infected cells. However, if there are other RNA species coded for by the viral strand, its abundance should be less than 15% of this species. Similar analysis of poly(A)[−] RNA from infected cells revealed only a minor fraction of short heterogeneous RNA–DNA hybrids, with no accumulation of any particular length.

It is thought that the MVM genome has a small number of structural genes, perhaps only one¹⁶. The coding capacities of

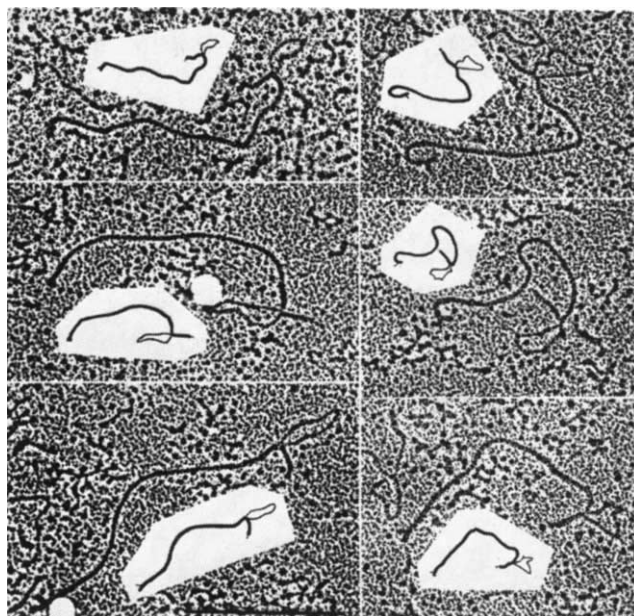


Fig. 2 Electron micrographs and their tracings of hybrids formed between cytoplasmic poly(A)⁺ RNA and MVM DNA. Cytoplasmic poly(A)⁺ RNA was prepared, annealed with genomic DNA and visualised in the electron microscope as detailed in Fig. 1. Scale bar, 0.5 µm.

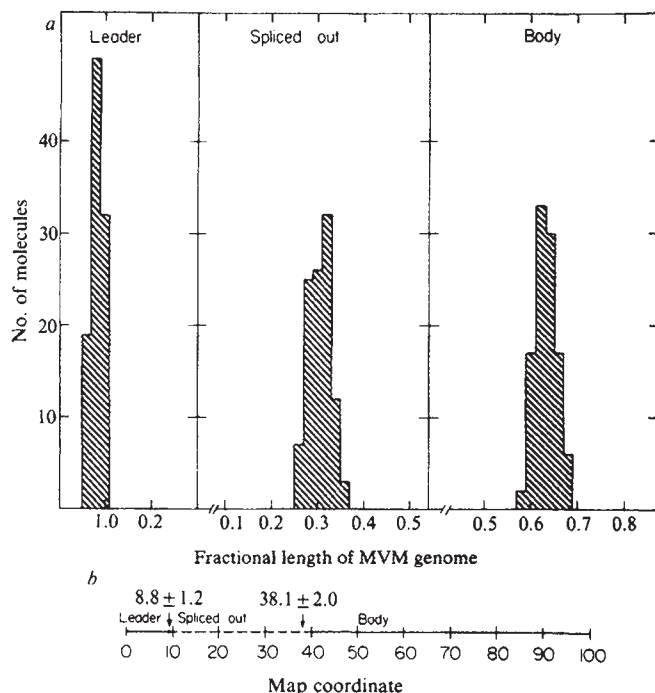


Fig. 3 *a*, Histograms of the leader, intron and body obtained from analysing structures as shown in Fig. 2. One map unit (MU) corresponds to 1% of the length of the entire structure. *b* Shows the coordinates of each part on the viral DNA. The 3'- and 5'-ends of the genomic DNA were taken as 0 and 100 MU, respectively. Leader, 8.8 ± 1.2 MU; spliced out, 29.3 ± 2.4 MU; body, 61.8 ± 2.3 MU.

the leader and body of the poly(A)⁺ RNA are about 12,000 and 90,000 daltons, respectively. If only the body codes for polypeptides, as in adenovirus¹ and SV40 late mRNAs², then it could code for polypeptide A (83,000 daltons) and protein B would be processed from this. If this is true then the viral mRNA would represent only about 70% of the viral genome and its primary transcript. Elucidation of the role of the remaining 30% of the primary transcript would be of great interest.

It is possible, however, that the MVM genome contains several genes, as a result of more than one functional initiation or termination site during translation. In this respect, it is noteworthy that splicing of RNA sequences at the post-transcriptional level could give rise to several functional mRNAs from the same precursor RNA, as has been found for SV40 early mRNAs¹⁷. These mRNAs could be translated in the same reading frame or switched from frame to frame.

The splicing of MVM RNA described in the present study resembles that of the mRNAs of other viruses and of eukaryotic cells but seems to be considerably less complex. Therefore, the transcription and processing of parvovirus mRNA provides a very useful model system for studying the mechanism of splicing and the physiological significance of this novel process in animal cells.

This research was supported in part by the Israeli Cancer Association and in part by the NCI. J.T. thanks the International Union Against Cancer for a travel fellowship.

JACOV TAL
DINA RON

Department of Biology,
Ben Gurion University of the Negev,
Beer Sheva, Israel

PETER TATTERSALL

Imperial Cancer Research Fund,
Mill Hill Laboratories, London NW7, UK

SUSAN BRATOSIN
YOSEF ALONI*

Department of Genetics,
The Weizmann Institute of Science,
Rehovot, Israel

Received 14 February; accepted 23 April 1979.

* To whom reprint requests should be addressed.

1. Berget, S. M., Moore, C. & Sharp, P. A. *Proc. natn. Acad. Sci. U.S.A.* **74**, 3171-3175 (1977).
2. Aloni, Y., Dhar, R., Laub, O., Horowitz, M. & Khoury, G. *Proc. natn. Acad. Sci. U.S.A.* **74**, 3686-3690 (1977).
3. Chow, L. T., Gelinas, R. E., Broker, T. R. & Roberts, R. J. *Cell* **12**, 1-8 (1977).
4. Klessig, D. F. *Cell* **12**, 9-21 (1977).
5. Horowitz, M., Bratosin, S. & Aloni, Y. *Nucleic Acids Res.* **5**, 4663-4675 (1978).
6. Panet, A., Gorecki, M., Bratosin, S. & Aloni, Y. *Nucleic Acids Res.* **5**, 3219-3230 (1978).
7. Rothenberg, E., Donoghue, D. J. & Baltimore, D. *Cell* **13**, 435-451 (1978).
8. Breathnach, R., Mandel, J. L. & Chambon, P. *Nature* **270**, 314-319 (1977).
9. Tilghman, S. M., Curtis, P. J., Tiemeier, D. C., Leder, P. & Weissmann, C. *Proc. natn. Acad. Sci. U.S.A.* **75**, 1309-1313 (1978).
10. Roop, D. R., Nordstrom, J. L., Tsai, S. Y., Tsai, M.-J. & O'Malley, B. W. *Cell* **15**, 671-685 (1978).
11. Tattersall, P. in *Replication of Mammalian Parvoviruses* (eds Ward, D. C. & Tattersall, P.) 53-72 (Cold Spring Harbor Press, 1978).
12. Bourguignon, G. J., Tattersall, P. & Ward, D. C. *J. Virol.* **20**, 290-306 (1976).
13. Chow, M. B. & Ward, D. C. in *Replication of Mammalian Parvoviruses* (eds Ward, D. C. & Tattersall, P.) 205-217 (Cold Spring Harbor Press, 1978).
14. Carter, B. J. in *Replication of Mammalian Parvoviruses* (eds Ward, D. C. & Tattersall, P.) 33-52 (Cold Spring Harbor Press, 1978).
15. Salzman, L. A. & Redler, B. *J. Virol.* **14**, 434-440 (1974).
16. Tattersall, P., Shatkin, A. J. & Ward, D. C. *J. molec. Biol.* **111**, 375-394 (1977).
17. Crawford, L. V. *et al. Proc. natn. Acad. Sci. U.S.A.* **75**, 117-121 (1978).
18. Penman, S. *J. molec. Biol.* **17**, 117-130 (1966).
19. Aloni, Y., Shani, M. & Reuveni, Y. *Proc. natn. Acad. Sci. U.S.A.* **72**, 2587-2591 (1975).
20. Horowitz, M., Laub, O., Bratosin, S. & Aloni, Y. *Nature* **257**, 558-559 (1978).
21. Canaan, D., Kahana, C., Mukamel, A. & Groner, Y. *Proc. natn. Acad. Sci. U.S.A.* (in the press).
22. Bratosin, S., Horowitz, M., Laub, O. & Aloni, Y. *Cell* **13**, 783-790 (1978).

RNA polymerase unwinds an 11-base pair segment of a phage T7 promoter

WHEN RNA polymerase binds to a promoter site, it must unwind part of the DNA double helix so as to expose the template bases. Wang *et al.*¹, Melnikova *et al.*² and Hsieh and Wang³ have used different experimental techniques to measure the degree of unwinding. Their estimates range from 7 (ref. 1) to 15 (ref. 2) base pairs unwound per bound *Escherichia coli* RNA polymerase (EC 2.7.7.6). These studies, however, do not establish which part of a promoter is melted out. A new technique described here directly proves unwinding and furthermore identifies the exact region that RNA polymerase opens in the A3 promoter of phage T7. This promoter is one of three strong *E. coli* RNA polymerase promoters used early in the life cycle of phage T7 (ref. 4).

Dimethyl sulphate (DMS) methylates the N-7 position of G and the N-3 of A on double-stranded DNA⁵. On single-stranded DNA, however, the N-1 of A and to a much lesser extent, the N-3 of C can also be methylated⁶; these positions are not accessible on double-stranded DNA because they are involved in hydrogen bonding (Fig. 1).

My technique takes advantage of this difference between single and double stranded DNA. First, I isolate the promoter on a restriction fragment which carries a ³²P label on the 5'-end of one strand only⁶. Then a complex between RNA polymerase holoenzyme and this promoter fragment is exposed to DMS. Wherever the RNA polymerase has unwound the double-stranded DNA, the N-1 position of A can be methylated. Precipitation of the DNA removes the RNA polymerase, allowing the two DNA strands to renature, except where the N-1 of A is methylated. The single strand-specific nuclease, S₁, which cuts at unpaired bases⁷, will cut the DNA strand at mismatched A

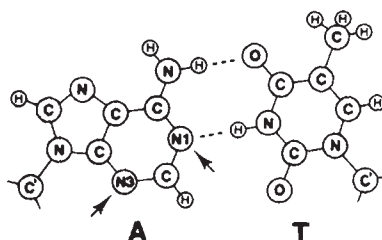


Fig. 1 AT base pair showing sites of methylation (see text).

and T bases created by methylation. This generates labelled strands, which terminate at the S_1 cutting sites. When the S_1 products are electrophoresed on a denaturing polyacrylamide gel⁶ next to the sequence of the labelled strand, the lengths of the labelled S_1 -cleaved strands identify the base pairs which were originally melted out by RNA polymerase.

Figure 2 shows an autoradiogram of such an experiment. The promoter fragment is an *HhaI*-*AluI* fragment of phage T7, carrying the A3 promoter, with transcription proceeding towards the *AluI* end⁴. As the *HhaI* end was labelled, Fig. 2 shows the S_1 cuts on the sense strand. Due to the 3'-overhang at

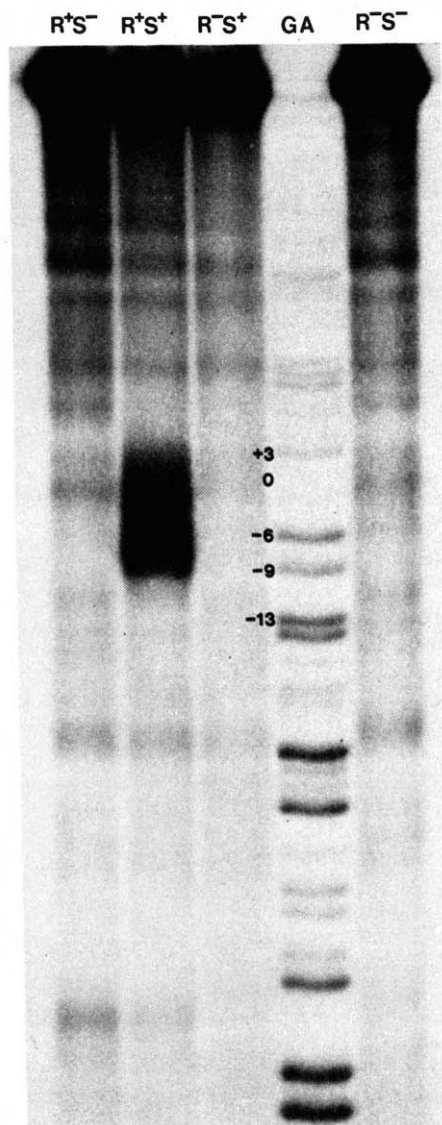


Fig. 2 Autoradiogram of a 12% polyacrylamide, 7 M urea gel⁶ with the S_1 nuclease digest of the A3 promoter fragment, which was previously methylated while complexed with RNA polymerase, in lane (R+S+). In the control lanes either S_1 (R+S-) or RNA polymerase (R-S+) or both S_1 and RNA polymerase (R-S-) were left out of the otherwise identical procedure. In another control, the DMS methylation was omitted from the protocol and again no S_1 cleavage was observed (not shown here). The G/A lane represents the sequence of the fragment according to the 'G greater than A' method of Maxam and Gilbert⁶. The numbers indicate positions of the promoter sequence (see Fig. 3). A 0.5–1- μ g sample of the 155-base pair long *HhaI*-*AluI* fragment was preincubated with 9–10 μ g of RNA polymerase holoenzyme (greater than 90% saturation with the σ subunit) in 100 μ l of binding buffer (0.01 M $MgCl_2$, 0.1 mM EDTA, 1 mM dithiothreitol, 50 mM Na cacodylate pH 7.3, 0.05 M KCl) for 1–2 min at 37 °C. Then 3–4 μ l of a 10.4 M DMS solution were slowly added during the reaction, which was quenched (see ref. 6) after 3 min by adding 100 μ l of 1 M mercaptoethanol, 1% SDS and 2 M NH_4 acetate, pH 5 (acidic conditions stabilise the labile N-1-methyl of A (ref. 5)). After precipitation with 2.5 volumes ethanol, the DNA was digested with approximately 4 units of S_1 nuclease in 50 μ l of S_1 buffer⁸ (0.03 M Na acetate, pH 4.6, 0.05 M NaCl, 1 mM $ZnSO_4$, 5% glycerol) for 1 min at 37 °C. The reaction was stopped with EDTA and SDS.

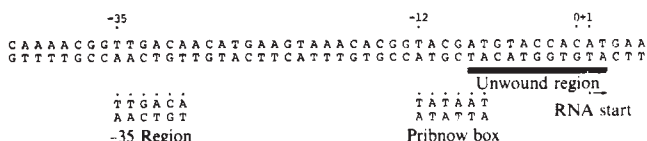


Fig. 3 A3 promoter sequence⁴. The canonical sequences⁴ of the Pribnow box and the -35 region homologies, the initiation site and the unwound region are shown.

the *HhaI* site, the 5'-label here should not be accessible to S_1 attack; nonetheless, about 25% of the counts cannot be precipitated after S_1 incubation, which may reflect 'breathing' of the ends.

The experiment in Fig. 2 visualises the unwound region in lane (R+S+) and shows it to be an 11-base pair segment extending from about the middle of the Pribnow box, a region of strong promoter homology⁴, to just past the initiation site (Fig. 3). The same limits appear in the analogous experiment with the antisense strand (labelled at the *AluI* end) (data not shown). This proves that there was no significant nibbling back by S_1 at the nicks, as that would shorten the size of the S_1 -cleaved strands and, therefore, shift the limits for unwinding towards the 5'-labels of either strand. The blurring of the bands is probably due to extensive methylation which will insert extra positive charges.

S_1 cuts appear at almost every position of the unwound region (except around the two successive CG base pairs). Most probably, the limited S_1 digestion causes cuts at either 5'- or 3'-side of a mismatched A or T, generating two labelled strand fragments which are separated by one nucleotide. This would account for S_1 -cleaved strands migrating with Cs or Gs of the sequence; no significant methylation at the N-3 of C need have occurred. (Due to the 3'-OH end of S_1 -cleaved strand fragments⁸, they will migrate slightly more slowly than the corresponding sequence strand fragment, which ends in 3'P (ref. 9).)

The unwound region I detect represents a minimum sequence, because if RNA polymerase were very tightly associated with the N-1 of an A: for example, it would prevent the methylation of that position and consequently the detection of such a melted base pair. Also, the CG base pairs bounding the unwound region may or may not have been opened.

The experiment directly shows that RNA polymerase unwinds a region covering 11 base pairs of the A3 promoter. This agrees well with previous estimates¹⁻³. The unwound region includes the initiation site and, more surprisingly, the latter half of the Pribnow box homology, suggesting that the Pribnow box may serve to initiate unwinding. I have probed for specific contacts between the A3 promoter and the polymerase with DMS and ethylnitrosourea according to techniques developed by Gilbert *et al.* (ref. 10 and unpublished data). Almost all contacts identified lie upstream from the unwound region (unpublished).

The new technique described here can detect any transiently opened AT base pairs; it should be useful in determining whether certain enzymes unwind DNA and, if so, which segments of the DNA they unwind.

I thank Walter Gilbert for advice and for help with the manuscript. This work was supported by an NIH grant (GM 09541-17) to Walter Gilbert.

ULRICH SIEBENLIST

Harvard Biological Laboratories,
16 Divinity Ave,
Cambridge, Massachusetts 02138

Received 5 February; accepted 24 April 1979.

- Wang, J. C., Jacobsen, J. H. & Saucier, J.-M. *Nucleic Acids Res.* **4**, 1225–1241 (1977).
- Melnikova, A., Beabekashvili, R. & Mirzabekov, A. D. *Eur. J. Biochem.* **84**, 301–309 (1978).
- Hsieh, T. & Wang, J. C. *Nucleic Acids Res.* **5**, 3337–3345 (1978).
- Siebenlist, U. *Nucleic Acids Res.* (in the press).
- Lawley, P. D. & Brookes, P. *Biochem. J.* **89**, 127–128 (1963).
- Maxam, A. & Gilbert, W. *Proc. natn. Acad. Sci. U.S.A.* **74**, 560–564 (1977).
- Shenk, T. E., Rhodes, C., Rigby, P. W. J. & Berg, P. *Proc. natn. Acad. Sci. U.S.A.* **72**, 989–993 (1975).
- Vogt, V. M. *Eur. J. Biochem.* **33**, 192–200 (1973).
- McConnell, D. J., Searcy, D. G. & Sutcliffe, J. G. *Nucleic Acids Res.* **5**, 1729–1739 (1978).
- Gilbert, W., Maxam, A. & Mirzabekov, A. in *Alfred Benzon Symp.* **9** (eds Kjeldgaard, N. O. & Maaløe, O.) 139–148 (Academic, New York, 1976).

reviews

Analytical history of ornament

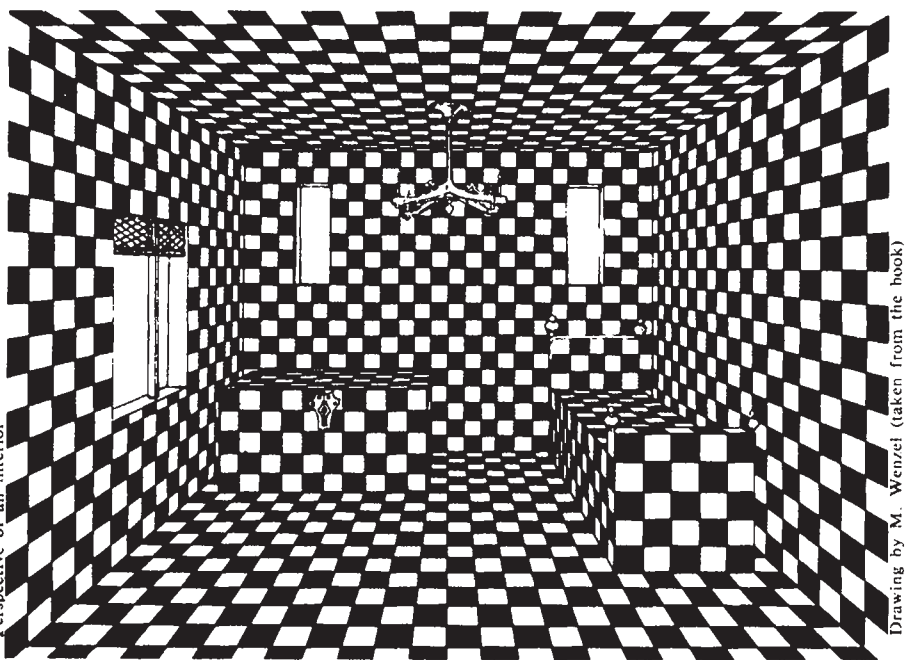
N. K. Humphrey

The Sense of Order: A Study in the Psychology of Decorative Art. By E. H. Gombrich. Pp. 411. (Phaidon: Oxford, 1979.) £15.

The Psychology of Decorative Art is the title of a book which has still to be written. That Gombrich should have used it as the sub-title of his latest essay in Art History is a pity for two reasons: first because it will mislead the reader as to what the book is really about; and second, because it has apparently misled the author into biting off more than he can chew. Despite its pretensions this book has very little of value to say about psychology. It has, however, a great deal to say about the more traditional subject matter of the art historian: craftsmanship, style, symbolism, tradition, patronage, and especially about what other critics have had to say about these things.

The body of the book is an analytical history of ornament. From time immemorial *homo faber*—man the serious-minded maker of useful artefacts—has openly cohabited with a more flippant mistress: where he might otherwise have been content with objects which were merely functional, she, his playful (and sometimes tiresome) muse, has persuaded him to decorate them. Her nature, it seems, abhors a visual vacuum; wherever she sees an empty surface, a hard edge, a simple corner, she yearns to overlay it or surround it with decorative patterns. So floors and walls are covered with geometrical designs; an arabesque runs around the cooking pot; the hilt of a sword is decked with jewels; stone columns are fluted and topped with a wreath of acanthus leaves; even the letters of the alphabet are adorned with knots and flourishes.

Gombrich brings to bear on these and related examples of human eccentricity his unrivalled combination of erudition, curiosity and taste. The first chapters deal with the criticism of ornament, the cult of purity and the debates about design in Victorian England. Next the craftsman's practice of pattern making is discussed as a response to the challenges of material, of geometrical laws and visual constraints. And, in perhaps the most intriguing sections of the book, the origins and fate of certain traditional motifs are followed through (the Paisley pattern,



Perspective of an interior

Drawing by M. Wenzel (taken from the book)

the acanthus scroll, and others). An Epilogue reviews some of the analogies between the spatial orders of decorative design and the temporal patterns of the dance, of poetry and, notably, of music.

Amidst all this there is, let it be said, a sprinkling of psychology, or what might less kindly be called 'psychologising'. Reference is made (particularly in three middle chapters) to "gestalt formation", "selective attention", "information theory", and so forth. But—and I do not think it is merely the closed-shop mentality of a professional psychologist which makes me say it—I found these excursions into experimental psychology for the most part trite, irrelevant or amateur. Although no-one has yet made a good job of linking art and psychology, Gombrich himself in his earlier book, *Art and Illusion*, has already made a much better job of it than he does here.

To the reader looking for a psychological explanation of design, however, the biggest disappointment will come with the vacuity of the 'biological' arguments which Gombrich propounds at the beginning of the book. It may be true that man and other animals have an innate 'sense of order', and that they find order relaxing and a mild

degree of disorder interesting (*ergo* pleasurable?). But Gombrich does not present a convincing case for why this should be so; nor, having set out the thesis, does he draw any but the most banal conclusion from it—that people like what we already know they like. For me at least, it adds nothing to the understanding of man's love of order to point out that "Nature around us is throbbing with complex rhythms, and these rhythms serve the purpose of life". True, man's heartbeat is rhythmical, and he would be dead if it were not; but, equally, his blood is red and he would be dead if it were not—and that surely does not explain why he prefers the colour red to blue.

The book is finely illustrated, and, as always, Gombrich proves a brilliant guide as to what to see and why. I shall never look at the roof of King's Chapel, or a William Morris wallpaper or a Paisley pattern tie in quite the same way again. But the way in which Gombrich succeeds in deepening our perception of these patterns is by telling us more about their makers rather than by telling us more about ourselves. □

N. K. Humphrey is Assistant Director of Research at the Sub-Department of Animal Behaviour, University of Cambridge, UK.

Disaster strikes from the blue

Peter J. Smith

Earthshock. By Basil Booth and Frank Fitch. Pp. 257. (Dent: London, Toronto and Melbourne; Walker: New York, 1979.) Hardback £6.95, \$11.95; paperback available from Sphere (London) in May, 1980.

WITHIN ten minutes of taking up this book I had come to two conclusions, both minor in themselves but both, it subsequently became clear, indicative of deeper malaise. The first was that someone has seen fit to impose a veneer of silly sensationalism on an already spectacular topic, presumably in the belief that a bit of spurious exaggeration does little harm to sales. It begins with the title. An earthshock, it seems, is simply a natural disaster. A short snappy title is all very well, and even to be applauded, but not at the expense of accuracy. As not all natural disasters (for example, glaciation) can be remotely regarded as "shocks", the authors are reduced to referring to "slow earthshocks", which reminds me of that superb character of childhood rhyme, the "bare-footed man with clogs on [who] came slowly running past".

Then there is the lurid dust jacket, showing a huge "tidal wave" about to overwhelm a Manhattan-style environment. Well, all right. More importantly, however, there is the jacket blurb which asks, "can the Earth survive?" and asserts that "A single great natural disaster . . . could destroy civilisation as we know it". For a start, no-one outside Monty Python refers any more to "civilisation as we know it", at least not without inviting ridicule. Yet the basic fault here is not some blurb-writer's exuberance but the authors, who adopt the same tone in parts of the text. They begin the book, for example, with a crude five-page account of the destruction (in 1987) of America's eastern seaboard by a fireball and its aftermath, and subsequently discuss nation-scale disaster from a serious scientific point of view. Is it worth spending time thinking about such unlikely (but admittedly possible) catastrophes when we cannot even cope adequately with events that are more familiar?

The second conclusion is that the book contains far too many flow-stoppers, ranging from the trivial to the more serious. The Earth's crust, mantle and core are referred to throughout somewhat eccentrically as

Crust, Mantle and Core. "Dilatancy" has become "dilatency". "Pangaea" appears as "Pangea"—a perfectly acceptable, indeed desirable, simplification in the context of the abolition of all ligatured diphthongs and their remnants, but presumably a mistake in a book that refers to aeolian, Palaeozoic and palaeomagnetism (and even, in one place, the idiosyncratic "palaeomagnetic"). Imperial units (feet, miles, and so on) are used throughout without even the parenthetical acknowledgement of metric units—a disgrace when (outside North America) a whole generation of children has been brought up on metric units. And even more seriously, one is continually being brought up sharply by the almost impenetrable: "However academic the study of plate movement may seem, its importance cannot be minimised". What does this mean?—the very opposite of what it actually says?

If that were the sole extent of the problem I could perhaps be accused of nit-picking. Unfortunately, however, the confusion goes much deeper, ruining some good ideas in the process. For example, before tackling natural disasters as such, Booth and Fitch devote some 50 or so pages to spelling out the nature of the generally more gradual processes involved in the Earth's behaviour. This is an excellent approach. All too often the authors of "popular" books on natural hazards give the impression that disaster strikes from the blue with little, if any, reference to the context in which terrestrial events and processes can combine to produce catastrophe. But to cover this background in so short a space is difficult and needs careful planning. Booth and Fitch do not entirely succeed. This section is rather confused, dense in concepts and terms all too frequently inadequately explained, and thus likely to be hard going for the non-expert.

The remainder of the book is devoted largely to particular natural hazards—to their general characteristics and to specific examples of disasters they have generated. Here the going is easier. To me, these chapters become more and more engrossing as the material becomes less and less familiar during the progression from earthquakes, through volcanoes, glaciation and flooding to extraterrestrial bombardment. But individual readers will have their own interests. To be objective about it, the book is probably most authoritative on those topics—especially volcanic matters—closest to the authors' hearts. On subjects closer to my heart, on the other hand, respect for the facts is not always all it should be. The notorious Denver earthquakes of the 1960s were not noticed by US

Geological Survey scientists in 1966 but by the alarmed inhabitants of Denver some years before; and the cause of the shocks was discovered by David Evans, a consultant geologist, in 1965. Moreover, the cause was not the injection into the ground of "industrial" waste fluids but of military waste material. The emotional overtones of the affair were thus rather different.

Some of the authors' judgements are also questionable. To say, as Booth and Fitch do, that "If the scientists involved in this work [that is, earthquake studies] are given sufficient encouragement and adequate financial support there is no reason why earthquake prediction should not become a routine matter within a decade" is really far too optimistic and therefore grossly misleading to the uninitiated. Indeed, it seems an unlikely prognosis even in respect of geologically comparatively simple areas such as the San Andreas fault zone of California. Moreover, to give the impression that prediction is purely or even primarily a scientific problem is dangerous. The social implications of a viable earthquake prediction capability are so complex that there must be real doubt whether non-totalitarian régimes could cope adequately with such a power without significant, and perhaps unacceptable, changes in socio-political attitudes.

Failure to give adequate attention to such factors is to demonstrate political naivety, a condition not uncommon among academic scientists. It is also naive to suggest, as Booth and Fitch do, that "the United Nations should have a single, comprehensive world natural disaster control agency with certain overriding powers, able to investigate and prepare for all types of natural disaster anywhere in the world, ready to go in with the right kind of advice and aid immediately it is required, without political let or hindrance". The United Nations has been conspicuously unsuccessful in almost everything it has ever tried to do precisely because such supranational authority is totally unacceptable to any nation of any consequence; and perhaps there is some merit in that uncooperative attitude when one takes account of the motives of the supranationalists. As Booth and Fitch put it, "people must be prevented from falling victim to their own or other people's stupidity and avarice by strictly enforced laws". Nasty.

Despite its descriptive merits, the long-awaited *Earthshock* is at best a disappointment and at worst a bit of a disaster. □

Peter J. Smith is Reader in Earth Sciences at the Open University, Milton Keynes, UK.

Travelling through time

Stephen Siklos

Time Warps. By John Gribbin. Pp. 167. (Dent: London, Melbourne and Toronto; Delacorte/Eleanor Friede: New York, 1979.) Hardback £5.95, \$8.95; paperback available from Sphere (London) in September, 1980.

Time Warps, the most recent addition to Dr Gribbin's list of popular science books, is about time and time-travel. It is broad in its scope; it begins with a discussion of time-telling through the ages, ends with an excursion into Taoist philosophy and the *I Ching*, and touches on Special Relativity, black holes, precognitive dreams and reincarnation in between.

For the most part, the author does not claim to offer any new insights; rather, as he says, to present old ones in a logically ordered structure. These ideas have been collected either from other popular science books, or from science fiction literature, and the book is liberally sprinkled with direct quotations from these sources. The author's own contribution consists of an explanation of déjà-vu experiences, all forms of dreams, and reincarnation in terms of travel "sideways in time". He postulates the simultaneous existence of all possible universes, and his theory is that dreams and similar phenomena arise from psychic contact with occurrences in some other universe, perhaps one very similar to our own. The beauty of this is that it explains not only dreams which come true, but also those which don't. Dr Gribbin may or may not be the first person to have thought of this idea, but as far as I know, he is the first to admit to it.

This book is aimed at a wide non-scientific audience. It is therefore particularly reprehensible that some of the most important scientific statements are wrong. For example, most of section II is based explicitly on the following incorrect argument. The steps are: (1) Einstein's Theory of General Relativity predicts that under some circumstances massive objects such as stars will collapse to form rotating black holes; (2) these black holes are described by the Kerr metric; and (3) one can therefore travel through a rotating black hole into a new universe. Statements (2) and (3) are false. The misunderstanding arises because the Kerr solution contains paths connecting the interior of a black hole with other regions of the solution, which could be

construed as new universes. Unfortunately, this solution is both vacuum (that is, it contains no matter) and stationary, so it does not represent a realistic collapse. Nor can it have a real observer moving in it (let alone a spaceship). Furthermore, perturbation analysis shows that solutions which resemble Kerr outside the horizon, but which could represent a realistic collapse and contain observers, cannot be extended through the inner horizon into a new universe. The point is that even infinitesimal deviations from the special symmetries of the Kerr solution (such as would be caused by the existence of a particle of matter) drastically alter the structure of the interior solution and destroy the supposed gateway to the new universe (see McNamara, *Proc. R. Soc.* **358**, 499; 1978). It is therefore generally believed that a rotating black hole singularity is not as Gribbin puts it "a different kettle of fish" from a non-rotating one—it is the same kettle. Travellers falling through the horizon could never escape and would eventually be crushed. Appropriately, the incorrect claims made in the book are accompanied by the wrong diagram: the Penrose diagram of a non-rotating charged black hole is shown.

The remaining passages about physics are, on the whole, accurate. These include good introductions to elementary quantum effects, and to the time dilation effects of Special and General Relativity, although even these are marred by a grotesque analogy between velocities and angles on a circle ($v < c$; $\theta < 360$), and also by the statement that time for an astronaut in closed orbit goes slower than time on Earth: it goes faster if the orbital radius is greater than one-and-a-half times the Earth's radius.

The passages not concerned with physics are characterised by their superficial and tendentious style. The description of Stonehenge as an astronomical computer, which opens section I, is typical. It is gleaned largely from the books by Hawkins (*Stonehenge Decoded*; Souvenir: London, 1967) and by Hoyle (*From Stonehenge to Modern Cosmology*; Freeman: San Francisco, 1972). No mention is made of the review by Professor R. J. C. Atkinson, entitled *Moonshine on Stonehenge* in which much of Hawkins' book is shown to be scientifically and archeologically unsound (*Antiquity*, 1966) or of the correspondence in the 1967 volume of *Antiquity*, where comments range from "dubious" to "untenable". The section then progresses to the subject of clocks, which is illustrated by a picture of a sundial, apparently capable of "all the accuracy

of a modern clock". It concludes with a discussion of time paradoxes and the meaning of time. The level of debate here, as typified by the subsection headed "Temporal Pigeon Holes and the Cosmic Postman", is most politely described as lightweight.

Clichés, polysyllabic humour, and other forms of turgid jocularity abound in this section. If you get past "Meanwhile, the poor man in the street was harried along, willy-nilly, into the 20th century", you are confronted with "intrepid circumterrestrial travellers", their "staid counterparts back home", the "change of date situation" and the fact that "even the common man and woman must come to grips with the subtleties of horological hairsplitting".

More serious, however, is the pernicious tendency to populate the book with 'goodies' (in whose work the author sees support for his theories) and 'baddies'. Thus the goodies are "respectable physicists, schooled in the best scientific tradition", "very solid members of the community", and "very sober scientists with impeccable academic credentials and years of research experience". On the other hand, we read that "Many theorists have an innate dislike for white holes, perhaps because they are only just coming to terms with the implications of black holes, and don't want to move too fast too soon"; that "archeologists cannot bring themselves to accept the subtleties which astronomers find in the Stonehenge computer"; that "there is not enough evidence yet to persuade the doubters" (for the existence of tachyons); and, incredibly, that Dr Gribbin sometimes wonders whether the "verbiage" in philosophical articles is not just "window dressing, to frighten off . . . people who are not 'philosophers'". Interested readers will find plenty more similar examples throughout the book.

Much the same can be said of section III, which purports to relate Eastern Philosophy to Western Science. To be fair, I think that this section will appeal to people with a taste for bizarre interpretations, and many readers will want to follow up the ideas in the bibliography provided. I cannot take these interpretations seriously, because they fail to explain the basic mechanisms and so seem to have no advantage over many less outlandish theories.

A glance at this review will reveal that *Timewarps* is likely to become a best-seller, and that many people will learn a lot of new ideas from it. It is a pity that it was not written with greater regard for accuracy. □

Stephen Siklos is a Research Assistant in Applied Mathematics at Queen Mary College, University of London, UK.

Sophisticated astronomy

Euan W. MacKie

Megaliths and Masterminds. By P. Lancaster Brown. Pp. 246. (Robert Hale: London; Scribners: New York 1979.) £5.50; \$14.95.

THE title suggests that here is another attempt to capitalise on the increasing interest, at many levels, in the evidence for sophisticated astronomical practices in the monuments of Neolithic and Early Bronze Age Europe. There must be at least a dozen books on the market at present with similar titles, and each with a picture of Stonehenge or some other megalithic site on the jacket, so the reviewer may be forgiven for having become somewhat hard to please in this area. These books vary considerably in quality, from the frankly mystical, through orthodox archaeological textbooks to excellent explanations of a difficult and complex subject (see *Nature*, 275, 75; 1978). One may well ask, what is the purpose of another one unless it includes original research or new ideas?

Mr Lancaster Brown in fact ranges more widely than the megaliths of the title, which he correctly defines as including both the standing stone sites of Britain and Brittany and the chambered burial mounds of Atlantic Europe from Spain to Scandinavia (although the latter also include dry-walled constructions and, if the collective burial rite is thought to be the decisive linking feature, some rock-cut tombs as well, which are in no sense megalithic). He has chapters on almost everything connected with astronomy in the ancient world, starting with Alexander Marshack's 'lunar notation' scratches on Palaeolithic bone implements. However, his suggestion that these marks, dating to well before 12,000 years ago, may be ancestral to the Irish Ogham script of less than 2,000 years ago will scarcely find favour with philologists and simply shows that speculating in the historical disciplines can seem too easy to the layman.

The author is, however, very interesting on the early development of research into archaeoastronomy and

ancient metrology, and this is perhaps the most useful part of the book. We are given vivid accounts of Sir Norman Lockyer, L. Piazzi Smyth and Sir Flinders Petrie at work among the temples, pyramids and obelisks of ancient Egypt and of the evolution of thought about astronomical qualities in British prehistoric sites from the seventeenth century onwards. It is very useful to be reminded that many of the modern ideas about ancient astronomy and metrology, developed in such detail by the Thoms, in fact derive from those of earlier workers, and salutary to learn again that some of these became obsessed with their theories to such an extent that they ignored or derided the historical and archaeological evidence. It is true that archaeologists were often of little help, usually being innately sceptical about archaeoastronomical ideas, but they were sometimes faced with such amazing nonsense that their scepticism is understandable.

The author provides one classic example from Lockyer's 1909 book on Stonehenge in which he discussed the living quarters of the astronomer-priests who, he assumed, were at work in Neolithic Britain. He supposed that damp was a severe problem so that shelters were needed, but assumed quite arbitrarily that the people of 4,000 years ago were entirely ignorant of carpentry. Thus they had to live in the megalithic chambered tombs of the period which were supposed to have been equipped with stone doors. (One would scarcely be surprised to read further that the doors often jammed, thus explaining the skeletons frequently found inside!) Even 70 years ago British archaeology had advanced far beyond that; presumably Lockyer had never entered a museum and seen a stone axe. Memories of this kind of thing still help to perpetuate a divide between workers like the Thoms and many orthodox archaeologists.

These historical insights are very useful but the chapters on the astronomy of the ancient Babylonians and of the American Indian civilisations are too brief to be of more than casual interest. Presumably the section on ley lines and 'geomancy' was included as another warning about the eccentric ideas that still flourish on the fringes of megalithic astronomy and which also help to keep many archaeologists at a distance. Perhaps this very breadth of coverage conceals another danger. The author often gives guidance and opinions on the present state of our knowledge of Neolithic Western Europe but these are based primarily on the technical archaeoastronomical evidence, itself only a small fragment of the total knowledge we have of the period concerned.

For example, there is not one but two vitally important aspects of hypotheses about the 'masterminds' among the megalith builders. The first of course concerns the plausibility of the sites claimed as observatories, or as showing high meteorological skill in measuring, and has been discussed extensively many times; this book deals mainly with it. However, the question of what kind of social organisation the people concerned had is an equally important but much trickier subject; it is rarely touched on by archaeoastronomers and their followers as it depends mainly on more orthodox archaeological evidence as well as on anthropological knowledge. There is no problem in ancient Egypt and Babylonia, or in Mesoamerica, which were urban societies known to have had specialist classes of priests and wise men, but what about Neolithic Europe? The population there was in a pre-urban stage and the usual view is that it was much more simply organised, not stratified into classes, and that it certainly did not possess élite orders of astronomer-priests and wise men of the kind which should have existed even if only half of what Professor Thom claims is correct.

The problem is whether the most important Neolithic sites and artefacts can be reinterpreted in terms of such a stratified, theocratic society. If none can, then the case for sophisticated ancient astronomy becomes much harder to maintain, no matter how impressive the technical evidence. The reviewer recently devoted half of an entire book (*Science and Society in Prehistoric Britain*, 1977) to raising this problem, and to making a start on resolving it, but feels he still has some way to go before convincing all his colleagues. Mr Lancaster Brown, however, only brings up the problem in a few lines in the final chapter, saying simply that we can be sure that such an élite class existed and that "Archaeological field evidence strongly supports this idea". Neither the book cited, nor Aubrey Burl's fundamental work *The Stone Circles of the British Isles* (1976) is mentioned anywhere and one must regretfully conclude that this is a further example of the idea that it is enough to be familiar mainly with the specialised archaeoastronomical evidence. Yet that is only the start of the journey, providing a new and clear window on to the past in addition to many older ones, by means of which our very complex but fragmentary picture of life in Neolithic Britain and Europe may eventually be substantially modified. □

Euan W. MacKie is Assistant Keeper in Archaeology and Anthropology at the Hunterian Museum, University of Glasgow, UK.

● In the review of *Physical Chemistry: Principles and Applications for the Biological Sciences* by Tinoco, Sauer and Wang, (*Nature*, 278, 85; 1979), it was incorrectly stated that a paperback edition had been published. A hardback edition only is available from Prentice-Hall International at £12.95, with a paperback *Solutions Manual* at £2.55.

Biological regulation processes

Biological Regulation and Development. Vol. 1: Gene Expression. Edited by R. F. Goldberger. Pp. 558. (Plenum: New York and London, 1979.) £24.88.

THE motivation to produce this inaugural volume of a multi-author treatise on regulation is that it would be productive to approach the subject in a way that cuts across traditional boundaries. The editors aim to avoid bringing together all possible facts relevant to a particular operon, virus or biosynthetic system. For instance, no one person works on suppression *per se*, only on aspects of the phenomenon. The actual chapter by Steege and Söll on suppression successfully picks up the various threads from many different laboratories. This approach, however, means that the same system is often discussed from different points of view by different authors. In a few places this needs some stronger editorial control and cross reference to assist readers with less experience in these areas. Dia-

grams of the same system sometimes appear quite differently in consecutive chapters. The chapter on regulation of DNA replication by Lark suffers from a complete lack of diagrams. In other chapters there is some obvious duplication of material which can be irritating. On balance, however, these problems do not greatly detract from what turns out to be a stimulating set of contributions stressing the essential concepts that underline our knowledge of regulation of gene expression. The book will be of considerable value to senior undergraduates, to postgraduates and to their teachers. Individual chapters deal mainly with regulation in microorganisms but recurrent themes are evolution of regulatory mechanisms and to some extent the relationships between prokaryotes and higher organisms. Later volumes must expand on this latter aspect but perversely no indication is provided, at least in this volume, of future contents.

The initial chapter by Goldberger covers general strategies of genetic regulation and provides a frame of reference for the discussions of operon complexity by Campbell and autogenous regulation by Savageau. Clarke develops the evolution theme with the

view that regulatory systems can be potent agents of evolutionary change. The chapters that follow deal with the relevance of DNA sequence and conformation. Pribnow's chapter on control signals should really be followed by the two more physicochemical contributions from Von Hippel and Sobell on the specificity of DNA-protein interactions and implications of DNA conformation. These latter chapters, careful consideration. Steitz then though speculative in places, deserve stresses the importance of RNA-RNA interactions, especially in ribosome function and proposes that mRNA effectively selects protein molecules that facilitate such interactions. Cortese follows with a review of the various functions of tRNA that are secondary to, or derived from, its primary function as amino acid adaptor. Finally, Maaløe amply demonstrates the stimulating but complex challenge of evaluating the possible interplay of all these various regulatory mechanisms in the growing organism (mainly *E. coli*).

Roy H. Burdon

Roy H. Burdon is Professor (Titular) of Biochemistry at the University of Glasgow, UK.

Thermal neutron scattering

Introduction to the Theory of Thermal Neutron Scattering. By G. L. Squires. Pp. 260. (Cambridge University Press: Cambridge, London, New York and Melbourne, 1978.) £16.

THERMAL neutrons are used widely to study structures and correlations in solids and liquids. They have just the right wavelength for atomic physics; and, as nuclear scattering takes place effectively at a point, one can use the scattered intensity to disentangle the motion of atoms. The momentum transfer dependence of the scattering is the Fourier transform of the spatial structure, and the energy dependence is likewise the Fourier transform of the time dependence of correlations. The magnetic moment of the neutron is a valuable bonus, sensitive to the magnetic structure of materials at the atomic level.

There are several textbooks on the physics which emerges, and on the techniques, but most authors content themselves with rough and ready derivations (or simply quoting) the rather complicated expression for scattering cross sections which are needed for quantitative analysis. This can be hard on research students wanting to extrapolate formulae to fresh situations and

wanting to grasp from first principles the meaning of phonons, magnons, spin waves and correlation functions. Gordon Squires' book fills this gap admirably. As the title implies, the book is aimed at deriving the formulae systematically and rigorously. It is the Goldberger and Watson of thermal neutron physics. Students will appreciate the beautifully precise, lucid and explicit algebra, which follows smoothly from undergraduate quantum mechanics. The material is extraordinarily well organised; there is great care over nomenclature, no mistakes and no missing lines. On the other hand, the casual reader might find the subject a bore, and even the dedicated reader will have to reflect carefully on the objectives of the formulae and the physical implications of each factor or approximation. To help his understanding, there are several problems, many of them quite hard, with solutions. Formulae are illustrated briefly with experimental results, but there is no systematic attempt to explore the physics which emerges from thermal neutron scattering.

This volume will be a valuable prop to the experimenter wanting to be sure of the foundation of his subject.

D. V. Bugg

D. V. Bugg is an SRC Fellow in the Department of Physics at Queen Mary College, University of London, UK.

SCIENTIFIC BOOKSHOP

H. K. LEWIS can supply works in all branches of Pure and Applied Science. Catalogues on request. Please state interests.

SCIENTIFIC LENDING LIBRARY

Annual Subscription from £7.50. (Available in U.K. only)

Reduced rates for multiple subscriptions.

Prospectus post free on request.

Quarterly List of New Books and new editions added to the Library sent post free to subscribers regularly.

H. K. LEWIS & Co. Ltd.

136 GOWER STREET,
LONDON, WC1E 6BS

Telephone: 01-387 4282
Telegrams: "Publicavl",
London, WC1E 6BS.

Circle No. 40 on Reader Enquiry Card.

Microbial life in extreme environments

Thermophilic Microorganisms and Life at High Temperatures. By T. D. Brock. Pp. 465. (Springer: New York, Heidelberg and Berlin, 1978.) \$29.70; DM54.

THIS book collates the ecological studies Professor Brock initiated and directed on microbial life in the hot springs of Yellowstone, some 1,200 miles from his University campus and at an altitude where water boils at 92 °C. The scope of the project was admittedly limited initially, but in the light of discoveries of unknown microbial species adapted to live in a variety of extreme environments, research ramified to include man-made thermal habitats and hot springs on a worldwide scale.

The objective was to study the structure, biochemical and growth characteristics, evolution and dispersal of such bacteria, blue-green algae and eukaryotic algae which seem to set the limits for life with respect to pH and temperature. Such natural model systems, with the relative constancy of their features, would then make possible deductions about long term ecological consequences of man-produced thermal and chemical pollution of the environment.

The cover picture showing the run-off channel of a Yellowstone hot spring is more suggestive of the scope of the book than its title. *Exploring Microbial Life in Hot Springs* might be a more appropriate title, which better expresses the 'adventure' character of the book.

The book stamps a record of personal achievement by drawing a demarcation line from earlier or contemporary work; it aims at helping future researchers, especially in Yellowstone, with the detailed description of the sites studied (chemical, physical and topographical). Furthermore, it includes valuable unpublished observations, suggests areas for future research and provides a wealth of ideas for *in situ* research approaches and improvised techniques that proved useful where conventional methods could not work or were unavailable. The author's confessions of being often misdirected or 'fooled', for example, by *Thermoplasma*, when in fact a new bacterium, *Sulfolobus*, had been isolated, illustrates the complexity and very often deceptive nature of working in an area where no precedent existed.

The seminar style of the book gives the reader both pleasure and a good share of the fascination and excitement

of the field work. The personalised manner in which the book was written is highlighted by the author's comments on his original paper on the genus *Sulfolobus*, that was rejected twice by the *Journal of Bacteriology*; and by the final chapter "Some Personal History" that includes a family and research team photograph.

The book illustrates effectively how research orientation can become usefully multi-directional in the way discoveries developed in a multi-disciplinary fashion, thereby becoming of interest to ecologists, environmental scientists, biochemists and geologists

alike. The author did not even miss the opportunity to comment on his observations on the effects of pH limits on the existence of higher organisms, including insects, frogs, birds, fish and plants.

Finally, this book, if read from cover to cover rather than used as a reference source, has all the elements that make it useful to ecologists. It would also make inspiring introductory reading for new researchers in microbial ecology. **G. D. Anagnostopoulos**

G. D. Anagnostopoulos is Lecturer in Microbiology at Queen Elizabeth College, University of London, UK.

Liquid structure in relation to their crystalline solids

The Molten State of Matter: Melting and Crystal Structure. By A. R. Ubbelohde. Pp. 454. (Wiley: Chichester, UK, and New York, 1979.) £23.50.

THIS book may mark the end of an era: or it may not. Scientific eras overlap, and a severe decimation of the time scale may be needed before the transition between them is marked with any definiteness. Knowledge of the figure of the Earth did not much assist the search for the sources of the Nile, and the Nile is still imperceptible on the geoid. This book is about liquids, and Ubbelohde has chosen its title to emphasise the fact that he will consider their structure in relation to that of their crystalline solids. Thus, attention is focused on phase change, and the relationship between the phases between which the change occurs: but one will find no mention here of the renormalisation group, or the names of Peierls, Landau, Wilson or Kadanoff, nor even, rather surprisingly, of Bernal and Fowler: which is not to say that there is not a wealth of references here to work in the present decade as well as past decades. Conversely, one may read a book about the renormalisation group and find mention of only three or four particular substances, if that. Ubbelohde deals with particular substances, and many comparative tables of the melting parameters of particular classes of substance are a feature of the book. These are instructive, and would be more so if one knew how much confidence to place in them. They seldom contain estimates of error, and sometimes fail to indicate source: there are discrepancies, sometimes minor and sometimes substantial, between values for

the same physical quantity in different tables: for example, the entropy of fusion of NaCl is 6.7 e.u. in Table 8.1 and 6.23 e.u. in Table 8.34. There are quite alarming discrepancies between heat capacity data for solid and liquid alkali metals in Table 9.1 and data for sodium in Fig. 9.1. Contradictory information about the magnetic susceptibility of iron phases at high temperature is given in Fig. 9.7 (p263) and the text on p264.

On p103 the reader is told, as though he should believe it, that for a second-order phase change two free energy surfaces must touch and intersect, with never a mention of bifurcation. Naïve readers must be warned against chapter 14, on "Liquid Crystals", shot through with misconceptions which should never have survived the classic paper of G. Friedel (1922).

Pleas for more observational data are a recurring refrain, and often justified: but, although on p90 the author complains about the lack of information regarding domain formation in order-disorder transformation, he never mentions antiphase boundaries, a keyphrase which could have led him to a wealth of material in the metallurgical literature.

It becomes apparent that several distinct and separate cultures grow on the one substrate of phase transitions. Of these, this book provides a useful literature guide to a substantial part of the physicochemical culture. Those readers who will take the trouble to check and amend the data will be in possession of a well-planned survey of empirical melting parameters, and some liquid properties for various classes of material. Considering the price to be charged, this should have been attended to by the publisher's editor.

F. C. Frank

Sir Charles Frank is Emeritus Professor of Physics at the University of Bristol, UK.

obituary

D. A. Hems

TO BE a successful scientist requires three attributes: intelligence, hard-work and perspicacity. Intelligence and hard work are not enough; perspicacity, insight, flair—call it what you will—make up a somewhat undefinable third requisite. Dr D. A. Hems, Reader in Biochemistry at St George's Hospital Medical School, London, who died tragically on 2 February 1979, possessed all three qualities. He was just 40, at the peak of his scientific career.

Hems received his schooling at the John Lyon school in Harrow, and entered University College, London as Bucknill Open Scholar in Medicine. At the end of his pre-clinical years, he took an intercalated Hons BSc in Physiology; during his clinical training at UCH he received prizes in neurology, and in obstetrics and gynaecology. On qualifying, he decided to pursue a scientific career, and moved to the Institute of Psychiatry to work on biochemical aspects of brain metabolism under Dr (now Prof.) R. Rodnight; three years later he received his PhD. Intelligence and hard work were clearly two of Hems' attributes.

Hems' interest now lay firmly in metabolism. But not in the elucidation of intermediates or in the properties of isolated enzymes. He realised that this aspect of biochemistry was, by 1966, proceeding on fairly predictable lines. What was lacking, was a proper correlation between the physiological behaviour of an organ and the biochemical activity of its cells. In Oxford, Sir Hans Krebs was developing the technique of whole organ perfusion. So it was to Oxford, on an MRC Research Fellowship, that Hems went in 1966: the first sign of that third attribute.

In 1969 Hems was appointed Lecturer in Biochemistry at Imperial College, under Professor Sir Ernst Chain. He now began to develop what proved to be his major scientific interest: the hormonal control of metabolism, using intact organs and cells. It was clear to him, as to others at this time, that the concept of a specific target organ for a specific hormone was too restrictive. He showed, for example, that the antidiuretic hormone vasopressin also has a direct action on the liver; that action is to stimulate glycogen phosphorylase, but not, as Hems discovered, by way of

cyclic AMP. Another of Hems' interests lay in obesity. With characteristic insight he chose to investigate fat synthesis in liver, rather than in adipose tissue: a choice that is beginning to appear well-founded. More signs of that essential third quality; sadly, fulfillment of Hems' potential was not to be realised.

During his time at Imperial College, Hems established close links with the research being carried out by Dr Anne Beloff-Chain (Lady Chain). He also built up an active group of his own; when in 1976 he was offered a Senior Lectureship in the new Biochemistry Department at St George's, the entire group chose to move with him to Tooting. Hems—Doug to all who were close to him—was at once immensely kind and considerate, helpful and inspiring, and it is easy to see why he was able to attract bright and energetic students, as well as more senior research workers. As a teacher, also, Hems was dedicated to his students.

His scientific achievements led to his appointment as an editor, first for *Clinical Science and Molecular Medicine*, and next of the *Biochemical Journal*. He also played an active part in the Regulation in Metabolism Group of the Biochemical Society.

Hems loved music and poetry; he particularly enjoyed witty and pithy comments on life, and was himself gifted with a dry sense of humour. The latter contributed to his demand as a guest speaker: at the time of his death, he had invitations to lecture at Guildford, London, Cairo, Cambridge and Copenhagen. But however much his scientific colleagues will miss him, it is his widow, Phyl Hems, and his two children Jacky and Clare, who face a sad and lonely future without Doug. It is to them that our sympathy is extended.

C. A. Pasternak

Ralph Emerson

WITH the death on 12 March 1979 of Professor Ralph Emerson, mycology has lost one of its brightest stars. He was born in New York City in 1912. He studied biology at Harvard obtaining his BS in 1933 and his PhD in 1937. He then came to England with a National Research Council fellowship and worked in the botany school at Cambridge for the next two years. In 1940 he joined the faculty of the University of Cali-

fornia at Berkeley becoming a full professor in 1953 and being chairman of the botany department from 1967 to 1971.

He was elected to the National Academy of Sciences in 1970. He served as president of the Botanical Society of America (1967), as president of the Mycological Society of America (1953) and as vice-president of the British Mycological Society (1971). In 1964 he received the Merit Award of the Botanical Society of America. To the First International Mycological Congress at Exeter in 1971 he contributed a brilliant general lecture which ended in a standing ovation.

His greatest scientific contributions were to the biology of water-moulds especially those belonging to the Blastocladales. Hans Kniep (1881–1930), that great student of sexuality in fungi, discovered in the last years of his life a unique sexual cycle in the water-mould *Allomyces*. Ralph Emerson developed the study of sex in species of that genus in beautiful detail. His 'An experimental study of the life-cycles and taxonomy of *Allomyces*' (*Lloydia* 4, 77–144; 1941) has become one of the classic papers of mycology. He was also responsible for a film 'Syngamy and alteration of generations in *Allomyces*', one of the finest mycological films ever produced. Further, Ralph Emerson made important contributions to the physiology of water-moulds. He found that certain species require an environment with a high concentration of carbon dioxide if resting spores are to be formed, whilst in one species oxygen must be absent if any growth is to occur, a most unusual situation in fungi.

He maintained that, although Blastocladales were clearly not organisms of economic importance, their study could make valuable contributions to fundamental problems. He inspired students and colleagues to undertake work that has fully vindicated this view. He launched Dr Cantino on his productive study of *Blastocladiella* which has done so much to illuminate the relationship between biochemistry and morphology. He inspired the late Dr Machlis and his colleagues to investigate the hormone concerned with the attraction of male to female gametes in *Allomyces*. This led in 1968 to the

resolution of the structure of sirenin, the first plant sex hormone to be characterised. All over America there are mycologists of distinction who received their early inspiration from Ralph Emerson.

During the war years 1944-46 he served as microbiologist to the Emergency Rubber Project in the United States which led him temporarily into the field of thermophilic organisms, particularly fungi responsible for development of high temperatures in compost heaps. After the war, although Blastocladales remained his main research area, he continued his war-time interest and with Dr Cooney was responsible for an important book *Thermophilic Fungi* published in 1964.

Ralph Emerson was much concerned with the basic causes of the student troubles on the Berkeley campus in the late sixties. In particular he felt it his duty to respond to his students' demands for relevance in mycological teaching. He immediately set to work to organise what he described as 'a running story of fungi and mankind with whatever minimum mycological technicalities had to be provided'. He devised a course that was both academically sound and excitingly relevant.

Ralph Emerson was a man of strong

character and great personal charm. Not only was he a research worker of outstanding ability, but he also wrote and spoke with elegance and clarity. His enthusiasm for his subject and for the experimental approach was unbounded. His former students and mycologists throughout the world mourn the loss of an eminent scientist. Above all he was a very special kind of person.

C. T. Ingold

David Zimmerman

It was with extreme regret that I heard of the sudden death of Dr David Zimmerman, who died on 10 November 1978 of haemorrhage following an injury to his eye by a tennis ball.

He came to England in the sixties with an MSc in physics from the University of Wisconsin and joined Dr Aitken in the Research Laboratory for Archaeology in the University of Oxford. Here he became one of the most productive research workers in the new field of thermoluminescent dating.

One of the factors which made absolute dating difficult was lack of homogeneity in the samples of ancient pottery to which it could best be applied. Dr Zimmerman developed a way of avoiding this problem by the separating out of grains in the sample which

were small compared with the range of alpha particles and in which therefore a statistical uniformity could be expected.

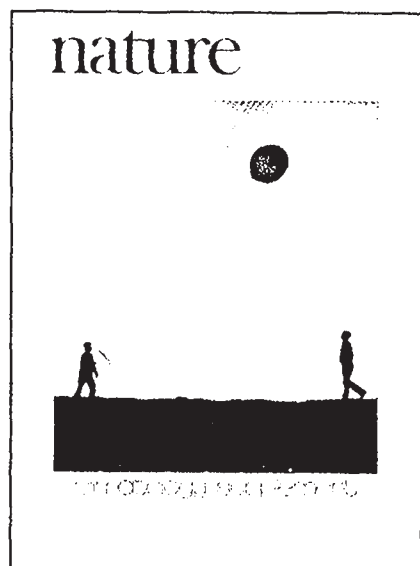
After gaining a well-deserved D.Phil as a result of this work, he returned to the United States where he continued to work on thermoluminescent dating of pottery at Washington University, St. Louis. Here he explored the possibility of using the zircon inclusions which usually contain so much uranium and thorium as to be almost entirely unaffected by background radiations, which are often impossible to estimate accurately.

Dr Zimmerman had many interests outside physics which he applied with characteristic success; for example he obtained a Blue at Oxford for badminton and played tennis with enthusiasm (Intercollegiate-Caltech); it was tragic that this should have led to his death. He enjoyed walking and camping and was interested in piano music. He was forty years old at the time of his death and has left a wife, Joan who was a contemporary research worker in 'TL' at Oxford. During the last twelve months he had started a valuable newsletter *Ancient TL*. His death is a very great loss to the field of thermoluminescence as well as to his large number of friends.

J. H. Fremlin

nature special supplement Climatology

(23 November 1978)



Contents

- BJ Mason
The World Climate Programme
- MJ Ingram, DJ Underhill & TML Wigley
Historical Climatology
- Valmore C LaMarche, Jr
Tree-ring evidence of past climatic variability
- GJ Shutts & JSA Green
Mechanisms and models of climatic change
- A Gilchrist
Numerical simulation of climate and climatic change
- BA Thrush
Recent developments in the atmospheric chemistry
- George L Siscoe
Solar-terrestrial influences on weather and climate
- CE Leith
Predictability of climate
- MK Miles
Predicting temperature trends in the Northern Hemisphere to the year 2000

NOW REPRINTED IN BOOKLET FORM

Back numbers also available

Parasitology Supplement (36pp)
22 June 1978

Science in China (3 reports by
Joseph Needham) 31 August 1978

Prices including postage:

UK £1.00
Rest of World US \$3.50 (surface)
US \$5.00 (air)

Payment may be made in any currency at the prevailing rate of exchange. Cheques should be made payable to *Nature*.

To: *Nature*, Macmillan Journals Ltd.,
Brunel Road, Houndmills, Basingstoke, Hampshire,
England RG21 2XS

Please send me special supplement issues
in the quantities shown:

Climatology ☐ *Parasitology* ☐
(booklet) *Science in China* ☐

I enclose

Name

Address

Registered No: 785998 England
Registered Office: 4 Little Essex Street,
London WC2R 3LF